11-19857

1907–1908

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

ORMOND STONE W. E. BYERLY F. S. WOODS
E. V. HUNTINGTON G. A. BLISS
J. K. WHITTEMORE

PUBLISHED UNDER THE AUSPICES OF HARVARD UNIVERSITY

SECOND SERIES Vol. 9

FOR SALE BY
THE PUBLICATION OFFICE OF HARVARD UNIVERSITY
2 University Hall, Cambridge, Mass., U. S. A.

London LONGMANS, GREEN & Co.
39 Paternoster Row

Leipzig: OTTO HARRASSOWITZ
Querstrasse 14

THE CALCULATION OF LOGARITHMS

BY JAMES K. WHITTEMORE

CONTENTS

1. Introduction. The use of logarithms is so general that it seems desirable to give here a short account of the history of their invention and of the best method for their computation. To understand these explanations is necessary a comprehension of the following definitions and elementary properties of logarithms:

If $a^x = N$, where a and N are positive numbers, x is called the "logarithm of N with the base a." We write the relation in the form

$$x = \log_a N.$$

Two bases are in common use, the base ten used in numerical work, and the "natural" base, usually denoted by e, used in the calculus. The value of e is the sum of the infinite series:

$$e = 1 + \frac{1}{1} + \frac{1}{2!} + \frac{1}{3!} + \cdots$$
$$= 2.71828$$

correct to five decimals. If M is any positive number,

$$\log_a N + \log_a M = \log_a NM.$$

If p is any real number,

$$\log_a N^p = p \log_a N.$$

From these equations it follows that

$$\log_a N = \frac{\log_b N}{\log_b a} = \frac{1}{\log_N a},$$

where b is any positive number. If then the logarithms with one base b, of all positive numbers are known, the logarithms of all positive numbers with any base a, may be found by division.

We shall hereafter denote $\log_{10} N$ by $\log N$, and $\log_e N$ by lN. With this notation we may write

$$\log N = \frac{lN}{l10} = \log e \cdot lN.$$

It is proved in the calculus that

$$(1) \qquad l(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \qquad \text{if } |x| < 1,$$

where the symbol $|x|$ means the absolute or numerical value of x. From (1) it follows that, if $|x|$ be less than one,

$$(1') \qquad \log_a (1+x) = \log_a e \left\{ x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \right\}.$$

The factor, $\log_a e$, is called the modulus of the system of logarithms with the base a. We shall denote by M the modulus of the system of base ten. The value of M, correct to five decimals, is

$$M = \log e = 0.43429.$$

The series (1) is not convenient for the computation of logarithms, first because it does not converge if $|x|$ is greater than one, secondly because if $|x|$ is only a little less than one a very large number of terms must be taken to obtain an approximately correct value for $l(1+x)$. But the series is of

great importance in any discussion of the properties of logarithms, and moreover serves as a basis for the deduction of other series of great use in computation. One of the most useful of such series is the following:

$$
(2) \qquad \begin{aligned} ly &= \tfrac{1}{2}\, l(y+1) + \tfrac{1}{2}\, l(y-1) \\ &\quad + \frac{1}{2y^2-1} + \frac{1}{3}\left(\frac{1}{2y^2-1}\right)^3 + \frac{1}{5}\left(\frac{1}{2y^2-1}\right)^5 + \cdots. \end{aligned}
$$

This series converges for all values of y greater than one, and converges very rapidly for large values.

It may be deduced from (1) as follows: Replacing x by $-x$ in (1),

$$
l(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \cdots \qquad |x| < 1.
$$

Subtracting this series from (1) and dividing by two,

$$
\frac{1}{2}\, l\, \frac{1+x}{1-x} = x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots.
$$

Let x be written

$$
x = \frac{1}{2y^2-1}.
$$

Then in order that $|x|$ be less than one it is both necessary and sufficient that y^2 be greater than one. We find

$$
\begin{aligned} \frac{1}{2}\, l\, \frac{1+x}{1-x} = \frac{1}{2}\, l\, \frac{y^2}{y^2-1} &= ly - \frac{1}{2}\, l(y+1) - \frac{1}{2}\, l(y-1) \\ &= \frac{1}{2y^2-1} + \frac{1}{3}\left(\frac{1}{2y^2-1}\right)^3 + \frac{1}{5}\left(\frac{1}{2y^2-1}\right)^5 + \cdots, \end{aligned}
$$

from which series (2) is at once obtained.

For logarithms with the base ten, we have

$$
(2') \qquad \begin{aligned} \log y &= \tfrac{1}{2} \log (y+1) + \tfrac{1}{2} \log (y-1) \\ &\quad + M\left\{\frac{1}{2y^2-1} + \frac{1}{3}\left(\frac{1}{2y^2-1}\right)^3 + \frac{1}{5}\left(\frac{1}{2y^2-1}\right)^5 + \cdots\right\}. \end{aligned}
$$

If y is a prime integer greater than two, the numbers $y+1$ and $y-1$ are both even, and their logarithms may be expressed in terms of the logarithms of two and of integers less than y.

2. Historical. A detailed and interesting account of the invention of logarithms and of the methods of the first computers is given in the article on "logarithms" by J. W. L. Glaisher in the Encyclopedia Brittanica. From this I have taken most of the historical matter here given, and to this I refer the reader who desires a fuller account than is contained in this article.

Logarithms were invented, it is generally admitted, about 1614 by an Englishman, John Napier, Baron of Merchiston. His first published work was seen by Henry Briggs, at that time Professor of Geometry at Gresham College, London, later a professor at Oxford. Briggs visited Napier and worked with him towards the perfection of the theory, and subsequently devoted much time to the calculation of logarithms. From these two men the system of logarithms with the base ten has derived the name of Naperian or Briggsian logarithms. The word "logarithm," it may here be explained, comes from the Greek words, λόγων ἀριθμός, meaning the number of ratios, for logarithms were first regarded as a number of ratios. Thus if ten be regarded as the product of 10000 equal ratios, a, so that

$$a^{10000} = 10,$$

we may, since we find that approximately

$$a^{3010} = 2,$$

say that 0.3010 is the logarithm of two. Briggs was an ardent computer, and from 1614 to 1617 he calculated to fourteen decimals the logarithms of the integers from 1 to 1000. In 1624 he published his "Arithmetica Logarithmica" containing the logarithms to fourteen decimals of the integers from 1 to 20,000, and of those from 90,000 to 101,000. This earliest of all logarithmic tables contains all the logarithms necessary for our modern five- and six-place tables. Briggs was still occupied in calculating the logarithms of the integers between 20,000 and 90,000, when in 1628 Adrian Vlacq of Holland published a table giving to ten decimals the logarithms of all integers from 1 to 100,000. These tables have been the basis of compilation of nearly all tables of logarithms published since that time. Naturally they contained some errors.

It is noteworthy that these tables differ in arrangement from tables now in common use in that the logarithms are given to ten and fourteen decimals,

while the numbers, "arguments" of the table, contain only five figures. Modern four and five-place tables generally give the logarithm with only one more figure than is contained in the argument. That the latter arrangement is that best suited to practical numerical work appears from the following considerations: in such a table the difference of successive logarithms, that is, logarithms of arguments differing by one, is always less than forty-five units in the last decimal place, and for half of the values of the argument, less than ten units, so that interpolation for a logarithm of a number containing the same number of figures as the logarithm in the table may be performed mentally.* Now if two numbers are to be multiplied by the use of logarithms, and if at least one of these numbers is not known certainly beyond m significant figures, the product cannot be found correctly to more than m significant figures,* and just that number of figures will be given by the use of logarithms to m decimals.† Hence for multiplication of numbers of m figures it is desirable to use logarithms to m decimals, and consequently convenient, but not always necessary, to have a table of logarithms whose arguments contain $m - 1$ figures.

One may then reasonably inquire whether tables like those of Briggs and Vlacq have, in the last figures of the logarithms, any value. To this may be answered first that these tables were doubtless intended to serve as a basis for the subsequent calculation of larger tables; secondly that by interpolation the values of logarithms of intermediate numbers may be found from these tables to a large number of figures; indeed, as I shall prove,† the value of the logarithm to m decimals of any number may be found from an m-place table (one in which logarithms are given to m decimals) by interpolation with first differences if the argument is given with $\frac{1}{2}m + 1$ figures.‡

Briggs' method of calculating logarithms was extremely laborious, for at that time no developments in infinite series of the logarithm had been discovered. He extracted successive square roots of ten fifty-four times, obtaining the result,

$$10^{y} = 1 + 12781\ 91493\ 20032\ 35 \times 10^{-32} = 1 + a,$$

where
$$y = \frac{1}{2^{54}}.$$

---

* See Note A.

† See Note B.

‡ For a more exact statement, see Note B.

The right hand member of this equation may be written as one plus a decimal of thirty-two figures, fifteen ciphers preceeding the first significant figure of the fractional part. This result amounts to the statement that the Briggsian logarithm of the second member of this equation is

$$1/2^{54} = 55511\ 15123\ 12578\ 2702 \times 10^{-35}.$$

Briggs discovered that the logarithms of numbers of the form $1 + x$, where x is a decimal beginning with fifteen ciphers, are very nearly proportional to the decimal x. This ratio

$$\frac{\log (1 + x)}{x} = M,$$

approximately, for a small value of x, as we see from the series (1′). He obtained the value of the ratio by dividing $1/2^{54}$ by ten raised to that power, and obtained a result to eighteen decimals of which the first sixteen agree with the correct value of M. His process was to extract successive square roots of the number whose logarithm was sought, until a root of the form $1 + x$ was obtained. The logarithm of this root was then found by multiplying x by the ratio

$$\frac{\log (1 + a)}{a},$$

and from the logarithm of the root was found at once the logarithm of the number sought.*

Much interest was taken by mathematicians of the seventeenth and eighteenth centuries in the calculation of logarithms. Most computers used methods of an arithmetical nature similar to that invented by Briggs. But as the calculus developed, less painful methods based upon the use of infinite series were invented and used by some scholars, among whom was Newton. There have been few tables made from new calculations since the publication of Vlacq's work, though many writers have computed the logarithms of some numbers. Indeed a new computation of logarithms already known cannot be regarded as a very useful service to the mathematical community. The most important calculations, after those of Briggs and Vlacq, were made by two Englishmen, Sang and Thomson, and by direction of the French government for the "Tables du cadastre." Sang published in 1871 a seven-place table of the

* See Note C.

logarithms of the integers from 20,000 to 200,000, of which the last 100,000 were freshly computed by him. Thomson calculated to twelve figures the logarithms of all integers to 120,000. His work has been used to verify errors discovered in Vlacq's tables by a comparison with the "Tables du cadastre." An account of the latter tables may be of interest. A full description of them with explanations of the methods of their compilers is given by Lefort in an article in the fourth volume of the *Annales de l'Observatoire de Paris*.

In 1784 it was voted by the French authorities that new tables of the logarithms of numbers, of the trigonometric functions, and their logarithms should be calculated to correspond to the decimal division of the quadrant. The manuscripts of these tables, which have never been published, give to fourteen decimals the logarithms of the integers from 1 to 200,000, the natural sines, and the logarithms of sines and tangents. The intention was to publish them as twelve-place tables, but even the twelfth figure is not reliable. The work was done under the direction of an engineer, Prony. His subordinates were divided into three sections: first, five or six mathematicians, including Legendre, who were occupied with the preparation of formulas and with other purely analytical work; second, a group of seven or eight men who had at least enough mathematical knowledge to translate general formulas into numbers, and who did most of the calculating done from series; finally, a group of seventy or more computers who were occupied chiefly with the work of interpolation. The work was performed wholly in duplicate by independent workers, and required over two years for its accomplishment. The two manuscripts are deposited in Paris, one at the Observatory, the other at the Institute. The tables received, as has been stated, the name of the "Tables du cadastre." The exact meaning of the word, "cadastre" is not clear. In modern French usage it means a public register of the ownership of real property. The term is applied to other registers, especially to those carried out in great detail, and it is presumably in this sense that it was applied to these mathematical tables. The chief interest of these tables lies in the facts that no other computation on so great a scale has ever been carried through, and that this is by far the greatest work of interpolation ever undertaken.

The computation of the tables proceeded as follows. The logarithms of all integers from 1 to 10000 were computed to nineteen decimals, those of prime numbers from the series (2′), those of composite numbers by adding the logarithms of the factors. Thus was known the logarithm of every hundred from 1 to 1,000,000, and the logarithms of all integers between these

hundreds were obtained by interpolation. Only alternate logarithms were used as a basis for the interpolation, the others serving to check this work. The interpolation was not carried on by differences of known values, but the differences were themselves computed from series. The interpolation formulas used are of interest.

Let u_0, u_1, u_2, etc., represent successive values of the function considered. Thus,

$$u_0 = \log n, \qquad u_1 = \log(n+1), \text{ etc.}$$

We represent the first differences by Δu_0, Δu_1, etc., so that

$$u_1 = u_0 + \Delta u_1, \qquad u_2 = u_1 + \Delta u_1.$$

If the second differences are $\Delta^2 u_0$, $\Delta^2 u_1$ etc.,

$$\Delta u_1 = \Delta u_0 + \Delta^2 u_0, \qquad \Delta u_2 = \Delta u_1 + \Delta^2 u_1.$$

In general, if Δ^m be an mth difference

$$\Delta^{m-1} u_p = \Delta^{m-1} u_{p-1} + \Delta^m u_p.$$

By a combination of these formulas we find

$$u_p = u_0 + p\,\Delta u_0 + \frac{p(p-1)}{2}\,\Delta^2 u_0 + \frac{p(p-1)(p-2)}{3\,!}\,\Delta^3 u_0 + \cdots.$$

This formula is called by Lefort "Mouton's Interpolation Formula," but is an immediate consequence of Newton's formula.*

In this formula the differences may be calculated from series, for

$$\Delta u_0 = \log(n+1) - \log n = \log\left(1 + \frac{1}{n}\right)$$

$$= M\left(\frac{1}{n} - \frac{1}{3n^3} + \frac{1}{5n^5} - \cdots\right),$$

$$\Delta^2 u_0 = \Delta u_1 - \Delta u_0$$

$$= M\left(\frac{1}{n+1} - \frac{1}{3(n+1)^3} + \cdots\right) - M\left(\frac{1}{n} - \frac{1}{3n^3} + \cdots\right)$$

$$= -M\left(\frac{1}{n^2} - \frac{2}{n^3} + \cdots\right).$$

---

* Markoff, *Differenzenrechnung*, p. 15.

In this way may be obtained series for the differences of all orders. In the construction of the "Tables du cadastre" differences up to and including those of the sixth order were used. The tables, if completed, would have given the logarithms to twelve figures of all arguments with six figures, but as stated, the work was not carried beyond the finding of the logarithms of numbers up to 200,000.

3. Practical Computation. If a complete table of logarithms is to be calculated, there is doubtless no better method than that of interpolation, used in the Tables du cadastre. But if a single logarithm is to be calculated to a large number of decimals, the best method is one published by Weddle in 1845 in *The Mathematician*, a method which is not directly based on infinite series. Before examining this method it is interesting to know why it should ever be necessary to compute a logarithm to a large number of decimals. This will be made clear by the statement of Gernerth in the preface to his excellent five-place tables.* Every logarithm is given by him correct to five decimals, and moreover it is indicated whether this value is larger or smaller than the correct value. His table was compiled from Vega's Thesaurus, published in Leipzig in 1794, which gives logarithms to ten decimals. But Vega's last decimal place is sometimes wrong, and Gernerth decided to consider this tenth figure as unreliable. Finding in the Thesaurus

$$\log 5.873 = 0.768\ 8600\ 008,$$

Gernerth, fearing that the last figure might be nine units too large, was in doubt whether to write

$$\log 5.873 = 0.76886 +$$

or

$$\log 5.873 = 0.76886 -,$$

and was consequently obliged to compute the logarithm to ten places. It may be imagined that in some cases it would be necessary to compute a logarithm to even more than ten figures.

Again, for a single calculation, the easiest method may sometimes be to compute the logarithms needed for the work and finally to compute the number from the resulting logarithm. Thus, if it were required to find to thirty decimals the value of $1/23^{50}$, and only Gernerth's five-place tables were at our

* A. Gernerth, *Fünfstellige gemeine Logarithmen*, Vienna, 1901.

disposal, the best method would be to calculate to fifteen decimals log 23, and then from log $1/23^{50}$, to compute the number required.

In most practical work it never becomes necessary to compute a logarithm. It is seldom necessary to use logarithms to more than seven places, and generally four or five places suffice.

A description of Weddle's method, with an example of its application, is given by Gernerth, together with the small tables needed. This method I shall now describe and illustrate.

We suppose, as we may without loss of generality, the number N, whose logarithm is sought, to be between one and ten, and write

$$N = a + d,$$

where a is an integer less than ten, and d a decimal fraction.

We now divide $a + d$ by $a + 1$, writing the quotient, which is less than one, as a decimal carried to a number of places sufficient to insure the degree of accuracy desired.* This is next multiplied by $1 + c_r/10^r$, where c_r is the difference between nine and the first figure of the quotient not nine, which figure we suppose to be in the rth decimal place. Generally we should have $r = 1$. This product will have all figures, before the rth, nines and generally the rth also. This process of multiplication by factors of the form, $1 + c_r/10^r$, may be repeated until the final product is as near one as is desired.

Then, with as small an error as is desired

$$N = (a + 1) \div \left(1 + \frac{c_1}{10}\right)\left(1 + \frac{c_2}{10^2}\right)\cdots,$$

$$\log N = \log(a + 1) - \sum \log\left(1 + \frac{c_r}{10^r}\right).$$

To follow this method it is necessary to have a table giving the logarithms of integers from one to ten, and the logarithms of numbers of the form, $1 + c_r/10^r$, c_r being always an integer less than ten. Such a table, giving the logarithms to fifteen places for all values of r up to and including sixteen, is given by Gernerth on a single page.† It remains to explain how this short table may be computed.

---

* In the example, seventeen; see Note D.

† Loc. cit., p. 119.

Logarithms of numbers in the form $1 + \frac{c_r}{10^r}$ may be easily found from the series (1′), when the value of M is known. The logarithms of the prime digits 2, 3, 5, and 7, and the value of M are ingeniously found by J. C. Adams in the following manner:

Let
$$a = l\,\frac{10}{9} = -\,l\left(1 - \frac{1}{10}\right),$$
$$b = l\,\frac{25}{24} = -\,l\left(1 - \frac{4}{100}\right),$$
$$c = l\,\frac{81}{80} = \,l\left(1 + \frac{1}{80}\right),$$
$$d = l\,\frac{50}{49} = -\,l\left(1 - \frac{2}{100}\right),$$
$$e = l\,\frac{126}{125} = \,l\left(1 + \frac{8}{1000}\right).$$

The values of a, b, c, d, and e are easily obtained from the series (1). Then we shall have
$$l\,2 = 7a - 2b + 3c,$$
$$l\,3 = 11a - 3b + 5c,$$
$$l\,5 = 16a - 4b + 7c,$$
$$l\,7 = \tfrac{1}{2}(39a - 10b + 17c - d) = 19a - 4b + 8c + e.$$

The first of these formulas follows from the equation, easily verified,
$$2 = \left(\frac{10}{9}\right)^7 \left(\frac{24}{25}\right)^2 \left(\frac{81}{80}\right)^3.$$

The others may be similarly proved. From the two expressions for $l\,7$ we obtain a check formula
$$a - 2b + c = d + 2e.$$
We have further
$$\frac{1}{M} = l\,10 = l\,2 + l\,5 = 23a - 6b + 10c.$$

Adams used these formulas to find to 260 decimals the natural logarithms of 2, 3, 5, and 7, and the value of M to 282 decimals. It may be noticed that the discovery of such relations as those used by Adams would be a real application of simple indeterminate equations required to be solved in integers.

We proceed to calculate by Weddle's method the logarithm of 5873, using the short tables mentioned given by Gernerth. Since Gernerth's tables are carried to fifteen places, we shall proceed with the multiplication until the first eight figures of the product are nines, when the remaining seven factors necessary for the computation may be found by inspection. The work of division and multiplication must be carried to seventeen figures to insure the correctness of the fifteenth factor. To indicate multiplication by a factor $1 + c_r/10^r$, we place this number to the right of the multiplicand, write underneath the significant figures of the product by $c_r/10^r$, correct to seventeen decimals, and, adding this product to the multiplicand, obtain the product by $1 + c_r/10^r$. The factors obtained by inspection we shall denote by $1 + a_r/10^r$. We have

$$\log 5873 = 3 + \log 5.873$$

$$
\begin{array}{ccrc}
\dfrac{5.873}{6} & = & 0.97883\ 33333\ 33333\ 33 & 1+\dfrac{2}{10^2} \\
 & & 1957\ 66666\ 66666\ 67 & \\
\hline
 & & 0.99841\ 00000\ 00000\ 00 & 1+\dfrac{1}{10^3} \\
 & & 99\ 84100\ 00000\ 00 & \\
\hline
 & & 0.99940\ 84100\ 00000\ 00 & 1+\dfrac{5}{10^4} \\
 & & 49\ 97042\ 05000\ 00 & \\
\hline
 & & 0.99990\ 81142\ 05000\ 00 & 1+\dfrac{9}{10^5} \\
 & & 8\ 99917\ 30278\ 45 & \\
\hline
 & & 0.99999\ 81059\ 35278\ 45 & 1+\dfrac{1}{10^6} \\
 & & 9999\ 98105\ 94 & \\
\hline
 & & 0.99999\ 91059\ 33384\ 39 & 1+\dfrac{8}{10^7} \\
 & & 7999\ 99284\ 75 & \\
\hline
 & & 0.99999\ 99059\ 32669\ 14 & 1+\dfrac{9}{10^8} \\
 & & 899\ 99991\ 48 & \\
\hline
 & & 0.99999\ 99959\ 32660\ 62 &
\end{array}
$$

Then $c_1 = 0$, $c_2 = 2$, $c_3 = 1$, $c_4 = 5$, $c_5 = 9$, $c_6 = 1$, $c_7 = 8$, $c_8 = 9$. The remaining seven factors, $1 + a_r/10^r$, may be obtained by taking the differences

between nine and the successive figures to the fifteenth. The fifteenth figure is here correct because the errors of the approximate work cannot be greater than five in the seventeenth place.* Then

$$a_8 = 9,\ a_9 = 4,\ a_{10} = 0,\ a_{11} = 6,\ a_{12} = 7,\ a_{13} = 3,\ a_{14} = 3,\ a_{15} = 9.$$

There are two methods of using Gernerth's short table, one taking the logarithms as they are given, the other adding or subtracting to each logarithm, according as it is too small or too large, one fourth of a unit in the fifteenth place. The plus and minus signs written after the logarithms and the additional plus sign are used in the second method. The logarithms of the thirteen factors different from one are written in order, each correct to fifteen decimals.

$$\begin{array}{r l}
0.00860\ 01717\ 61918 & - \quad (25\cdot 10^{-17}) \\
0.00043\ 40774\ 79319 & - \\
0.00021\ 70929\ 72230 & + \\
0.00003\ 90847\ 44584 & + \\
04342\ 94265 & - \\
03474\ 35447 & - \\
00390\ 86502 & - \\
00017\ 37178 & - \\
26058 & - \\
03040 & + \\
00130 & + \\
00013 & + \\
4 & - \\
 & + \\
\hline
0.00929\ 12495\ 40788 & - \quad (50\cdot 10^{-17})
\end{array}$$

Now, $\log 6 = 0.77815\ 12503\ 83644\ -.$

Subtracting,

$$\log 6 - \sum_{r=1}^{8} \log\left(1 + \frac{c_r}{10^r}\right) - \sum_{r=9}^{15} \log\left(1 + \frac{a_r}{10^r}\right)$$

$$= \log 5.873 = 0.76886\ 00008\ 42856.$$

* Note D.

This value is liable to an error less than 7.55 units in the fifteenth decimal place if Gernerth's table be used by the first method; to an error less than 3.85 units in the fifteenth decimal place if the second method is used. It is only a coincidence that the two methods give the same result. The second method enables us to say that this value is correct to fourteen decimals, but not that the value of this correct to fourteen decimals,

$$0.76886\ 00008\ 4286,$$

is the true value correct to fourteen decimals. Either method enables us to say that the true value correct to fourteen decimals is either that just given or differs from that only in having the fourteenth figure a five.

4. The Limits of Error in Several Approximations. In preparing this sketch of the history of the development of logarithms and of the methods of their computation, I have frequently had occasion to question to what extent certain processes employed can be relied upon to give correct results. In attempting to answer the questions raised I have learned first, that we can in all cases assign to the error committed in the approximation a limit which cannot be surpassed, which I shall speak of as the limit of error; second, that these limits are often smaller than I had before supposed; and thirdly, in seeking to make the limit of error as small as possible, I found myself led in Weddle's method of computing a logarithm to very definite rules for computation. These results have interested me so much that I venture to hope that others may also find them interesting and not unprofitable. For a knowledge of the limit of error in computation will not only give to the computer more confidence in his result, but will also, it seems to me, give him a much more truly scientific habit of mind. I have divided these studies into four notes referred to in the previous pages. A few preliminaries will be useful.

It is customary in giving the approximate value x', of a number x, to give x' to a stated number of decimals m, so that $|x - x'|$ is as small as possible. Then x' is said to be the value of x "correct to m decimals." Clearly we shall always have

$$|x - x'| \leqq \frac{1}{2 \cdot 10^m}.$$

If x and y are any two numbers such that

$$|x - y| \leqq \frac{1}{2 \cdot 10^m},$$

we shall find it convenient to say that y is equal to x correct to m decimals. But we must note that if x' and y' are the values of x and y correct to m decimals, we may have

$$|x' - y'| = \frac{1}{10^m}.$$

A few examples will make these points clearer. Let $x = 0.43429$; then if m is three, $x' = 0.434$; if m is four $x' = 0.4343$; if $y = 0.4346$, y is equal to x correct to three decimals, but for m equal three, $y' = 0.435$ and $y' - x' = 1/10^3$.

Finally if x and y are two numbers such that

$$|x - y| \leqq \frac{1}{10^m}$$

then also

$$|x' - y'| \leqq \frac{1}{10^m}.$$

For example, taking again m as three, and $x = 0.43429$, if $|x - y| \leqq 1/10^3$, y lies between, or is equal to one of the values

$$0.43329, \qquad 0.43529,$$

and y' is equal to one of the values

$$0.433, \qquad 0.434, \qquad 0.435,$$

no one of which differs from x' by more than $1/10^3$.

We shall have occasion to use the developments for small values of x

$$\left.\begin{aligned} \log(1 + x) &= M\left\{x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \cdots\right\}, \\ -\log(1 - x) &= M\left\{x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} \cdots\right\}. \end{aligned}\right. \qquad 0 < x < 1,$$

From these developments we may infer that, for positive values of x less than one,

$$0 < \log(1 + x) < Mx,$$

$$0 < -\log(1 - x) < \frac{Mx}{1 - x}.$$

Note A. 1. If two numbers, of which one is not known beyond m significant figures are multiplied, the product cannot be known beyond m significant figures.

2. In an m-place logarithm table whose arguments contain $m - 1$ figures successive tabulated values differ by not more than forty-five units in the last decimal place, and for one half of the arguments by less than ten units.

Proof of 1.

We may, without loss of generality, suppose both of the numbers to lie between 1 and 0.1, so that each number has its first significant figure in the first decimal place. If the two numbers are x and y, and they are known to m and n figures respectively, and if the approximately correct values be x' and y', then

$$x' = x + \epsilon, \quad y' = y + \eta,$$

where

$$|\epsilon| \leqq \frac{1}{2\cdot 10^m}, \qquad |\eta| \leqq \frac{1}{2\cdot 10^n}, \qquad n \geqq m.$$

Then is

$$x'y' = xy + \epsilon y + \eta x + \epsilon\eta,$$

$$|x'y' - xy| = |\epsilon y + \eta x + \epsilon\eta|.$$

Since nothing is known of the signs of ϵ and η, we cannot know that the second member of the last equation is less than $|\epsilon y|$, which itself cannot be known to be less than or equal to $1/(2\cdot 10^{m+1})$, since y is greater than 0.1; hence it cannot be known that the value of $x'y'$ is that of xy correct to more than m figures. It may also be shown that the upper limit of $|x'y' - xy|$ is $(5.75)/10^m$, but that in more than two thirds of all possible cases this difference is less than $1/10^m$.

Proof of 2.

Let a and $a + 1$ be two successive values of the argument, and let the tabulated values of their logarithms, each correct to m decimals be $\log' a$ and $\log' (a + 1)$. Then

$$|\log a - \log' a| \leqq \frac{1}{2\cdot 10^m}, \quad |\log (a + 1) - \log'(a + 1)| \leqq \frac{1}{2\cdot 10^m}, \quad a \geqq 10^{m-2}.$$

The tabular difference, Δ, is

$$\Delta = \log'(a + 1) - \log' a = \log (a + 1) - \log a + \theta, \qquad |\theta| \leqq \frac{1}{10^m}.$$

Now is

$$\log(a+1) - \log a = \log\left(1+\frac{1}{a}\right) < \frac{M}{a} < \frac{44}{10^m}.$$

Then

$$\Delta < \frac{44}{10^m} + |\theta| < \frac{45}{10^m}.$$

If the first digit of a is five or greater, $\frac{M}{a}$ is less than $\frac{44}{5\cdot 10^m}$, and

$$\Delta < \frac{8.8}{10^m} + |\theta|$$

$$< \frac{9}{10^m}.$$

Note B. The Limit of Error in Interpolation by First Differences in a Table of Logarithms. Let us consider a table giving the logarithms correct to m decimals of all integers between 10^n and 10^{n+1}, a table then with arguments of $n+1$ figures. Let a be any value of the argument, and let $\log' a$ be the tabulated value of its logarithm. The formula for interpolation by first differences is

(3) $$\log(a+x) = \log' a + x\left[\log'(a+1) - \log' a\right],$$

where x is any positive number less than one. We proceed to investigate the limit of error of this formula. Let us write

(4) $$\log(a+x) = \log a + (x+\epsilon)\left[\log(a+1) - \log a\right].$$

We have from (1)

$$\log(a+x) - \log a = Ml\left(1+\frac{x}{a}\right)$$

$$= M\left[\frac{x}{a} - \frac{x^2}{2a^2} + \frac{x^3}{3a^3}\cdots\right],$$

$$\log(a+1) - \log a = M\left[\frac{1}{a} - \frac{1}{2a^2} + \frac{1}{3a^3}\cdots\right],$$

whence, from (4), cancelling $\frac{M}{a}$,

$$\epsilon\left[1 - \frac{1}{2a} + \frac{1}{3a^2}\cdots\right] = \frac{x(1-x)}{2a} - \frac{x(1-x^2)}{3a^2}\cdots$$

$$< \frac{x(1-x)}{2a},$$

since each term of the series in the second member of the last equation is less in absolute value than the preceding.* Moreover the coefficient of ϵ in the first member is greater than $1 - \frac{1}{2a}$. Hence we find

$$0 < \epsilon < \frac{x(1-x)}{2a-1}.$$

The maximum value of $x(1-x)$ is $\frac{1}{4}$. Then

$$0 < \epsilon < \frac{1}{4(2a-1)}.$$

Since we know that

$$0 < \log(a+1) - \log a < \frac{M}{a},$$

we may write

$$\log(a+x) = \log a + x\,[\log(a+1) - \log a] + \epsilon',$$

where

$$0 < \epsilon' < \frac{M}{4a(2a-1)}.$$

If we write

$$\log a = \log' a + \epsilon_1, \qquad \log(a+1) = \log'(a+1) + \epsilon_2$$

we know

$$|\epsilon_1| \leqq \frac{1}{2\cdot 10^m}, \qquad |\epsilon_2| \leqq \frac{1}{2\cdot 10^m}.$$

Now

$$\log(a+x) = \log' a + x\,[\log'(a+1) - \log' a] + \epsilon_1(1-x) + \epsilon_2 x + \epsilon'.$$

But

$$|\epsilon_1(1-x) + \epsilon_2 x| \leqq |\epsilon_1|\,(1-x) + |\epsilon_2|\,x \leqq \frac{1}{2\cdot 10^m}.$$

If now ϵ' is less than $\frac{1}{2\cdot 10^m}$, $\log(a+x)$ will differ from the value obtained from (3) by less than $1/10^m$. This will be the case if

$$\frac{M}{4a(2a-1)} < \frac{1}{2\cdot 10^m},$$

* The ratio of the nth term to the preceding is

$$\frac{1-x^n}{1-x^{n-1}} \cdot \frac{n}{n+1} \cdot \frac{1}{a} < \frac{n^2}{n^2-1} \cdot \frac{1}{a} < 1,$$

since $x < 1$ and $a > 10$.

or if

$$(4a - 1)^2 > 4 \cdot 10^m \cdot M + 1,$$

or finally, if

$$a > \frac{\sqrt{4 \cdot 10^m \cdot M + 1} + 1}{4}.$$

If m is even it may be easily seen that this inequality is satisfied by $a \geqq 10^{\frac{m}{2}}$; if m is odd, and equal to five or greater, it is satisfied by $a \geqq 1.05 \times 10^{\frac{m-1}{2}}$. Then in a table where m is even, if the number of figures in the argument, $n + 1$, is equal to $m/2 + 1$, the logarithm of any number may be found by interpolation with the formula (3) so that the error is not more than one in the last decimal place. Thus if m be four, there are needed for this degree of accuracy arguments of three figures. If m is odd and as great as five the number of figures needed in the argument to ensure this degree of accuracy is $(m + 1)/2$ if the first figures of a are as great as 105; for values of a beginning with figures less than 105, the arguments must be given with an additional figure.*

There results from these considerations this remarkable fact: a logarithm obtained by the formula (3) from m-place tables with arguments of $m/2 + 1$ or $(m + 1)/2$ figures,† when taken correct to m decimals, cannot differ by more than one unit in the mth decimal place from the true value of the logarithm correct to m decimals. This limit of error is the same that we should have if the process of interpolation were exact. For in that case, since each tabulated

* Gernerth, in his five-place tables, already referred to, uses arguments of four figures, with an additional table for arguments of five figures to 10800. This is, according to the theory here presented, one figure more in all cases than is necessary for reliable interpolation, but his arrangement evidently makes the interpolation much less laborious.

The seven-place tables of Dietrichkeit (Berlin, 1903) have arguments of four figures.

The eight-place tables published by direction of the French Ministry of War in 1891, and the eight-place tables of Mendizábal Tamborrel (Paris, 1891) both have arguments of five figures. According to this theory interpolation by first differences is reliable in these eight place tables, and also in the seven-place tables when a is as great as 1050.

The ten-place tables of Vega have arguments of five figures, so that interpolation by first differences is not reliable.

A brief treatment of the limit of error in a special case covered by my discussion is given by Markoff, *Differenzenrechnung*, pp. 35–39.

A rough discussion leading to an apparently similar result is given on page 63 of the tables of Dietrichkeit.

† According as m is even or odd, and when odd if $a > 1.05 \times 10^{\frac{m-1}{2}}$.

value is liable to an error of $1/(2\cdot 10^m)$, the value obtained for $\log(a + x)$ would be liable to an equal error, and consequently this value correct to m decimals might differ by one unit in the mth place from the true value correct to m decimals.*

We have, finally, supposing our m-place table to have arguments of $m/2 + 1$ or $(m + 1)/2$ figures,† to see with what limit of error a number N, may be found from its logarithm. We may, without restriction, suppose the characteristic of $\log N$ to be n, so that N lies between two successive arguments of the table, a and $a + 1$. Then a positive value of x, less than one, may be found from the equation

$$\log N = \log' a + x\,[\log'(a + 1) - \log' a].$$

The determination for N is then

$$N' = a + x.$$

Now the value determined by (3) for $\log N'$ is precisely $\log N$. Hence

$$|\log N - \log N'| < \frac{1}{2\cdot 10^m} + \frac{M}{4a(2a - 1)}.$$

If we write $N = A\cdot 10^{n+1}$ and $N' = A'\cdot 10^{n+1}$, A and A' both lie between 1 and 0.1, and

$$\frac{1}{M}\cdot|\log N - \log N'| = |lN - lN'| = l\left|\frac{N'}{N}\right| = l\left|\frac{A'}{A}\right|.$$

Let us suppose that A' is less than A. Should this not be the case the identical result would be obtained by interchanging these quantities in the following lines. We have

$$l\,\frac{A'}{A} = -\epsilon, \qquad \epsilon > 0.$$

Then is

$$A' = A\epsilon^{-\epsilon} = A\left(1 - \epsilon + \frac{\epsilon^2}{2!}\cdots\right),$$

* If in (3), $a + x$ is the value of a number N, correct to m figures, the limit of error in taking the second member of (3), correct to m places, for the value of $\log N$, correct to m places, is three in the last place even if a contains $m - 1$ figures. But if $a + x$ is the value of N, correct to $m + 1$ figures, the limit of error is, as before, one in the mth place, if $n = m/2$ when m is even, or if $n = (m - 1)/2$ and $a > 1.5 \times 10^n$ when m is odd and at least equal to five. Then in logarithmic work where numbers are only approximately known it is necessary, to obtain the best results, to use these numbers correct to one more figure than is contained in the tabulated logarithms.

† According as m is even or odd, and when odd if $a > 1.05 \times 10^{\frac{m-1}{2}}$.

and

$$|A' - A| < A\epsilon < A\left\{\frac{1}{2M.10^m} + \frac{1}{4a(2a-1)}\right\}.$$

Since $\frac{1}{2M} < 1.152$ and $4a(2a-1) > 7a^2$, we have

$$|A' - A| < A\left\{\frac{1.152}{10^m} + \frac{1}{7a^2}\right\}.$$

Now, however n is chosen, we cannot infer that $|A' - A| \leqq 1/10^m$, so that it will be impossible to assert that the values of A and A', each correct to m decimals, differ by less than two in the last place. But if n is chosen equal to $(m-1)/2$ or $m/2$, according as m is odd or even, it is easy to see by examining the different possibilities, that $|A' - A| < 2/10^m$, and consequently that the values of A and A', each correct to m decimals, cannot differ by more than two in the last place. The truth of this statement is immediately evident when m is even. If m is odd suppose, for example, that the first digit of a is 3; then is

$$|A' - A| < 0.4\left\{\frac{1.152}{10^m} + \frac{10}{63\cdot 10^m}\right\} < \frac{.53}{10^m}.$$

In fact it appears in this way that, if the first digit of a is seven or less, $|A' - A| < 1/10^m$. We may then say, returning to the numbers N and N', that the values of these two numbers each correct to m significant figures cannot differ by more than two in the last figure in any case, nor by more than one in the last figure if the first figure is less than eight.

Note C. The Limit of Error in Briggs' Method of Calculating a Logarithm. Briggs' method is based on the repeated extraction of square roots to a large number of decimals. It will perhaps be of interest to prove the fact, not always realized, that the square root of a number larger than one may be always found correct to at least as many decimals as the number is given, sometimes to more.

Let x' be the given value, correct to m decimals, of a number x. Then is $|x - x'| \leqq \frac{1}{2\cdot 10^m}$.

We seek the square root of x by finding the square root of x'. If this process is carried out correct to n decimals, and we write the resulting number

as $\sqrt{x''}$, we have

$$|\sqrt{x'} - \sqrt{x''}| \leqq \frac{1}{2\cdot 10^n}.$$

Now, the absolute value of the error,

$$|\sqrt{x} - \sqrt{x''}| \leqq |\sqrt{x} - \sqrt{x'}| + |\sqrt{x'} - \sqrt{x''}|$$

$$\leqq \frac{1}{2\cdot 10^m} \cdot \frac{1}{\sqrt{x} + \sqrt{x'}} + \frac{1}{2\cdot 10^n},$$

since

$$|\sqrt{x} - \sqrt{x'}| = \left|\frac{x - x'}{\sqrt{x} + \sqrt{x'}}\right| \leqq \frac{1}{2\cdot 10^m} \cdot \frac{1}{\sqrt{x} + \sqrt{x'}}.$$

Then if x is greater than one, we have

$$\sqrt{x} + \sqrt{x'} > 2, \qquad \text{and if } n = m + 1,$$

$$|\sqrt{x} - \sqrt{x''}| < \frac{1}{2\cdot 10^m}.$$

But the values of $\sqrt{x}$ and $\sqrt{x''}$, each correct to m decimals, may differ by one in the mth decimal place. We may notice that if $x > 10^{2p}$, we may, if we choose $n = m + p + 1$, assert

that
$$|\sqrt{x} - \sqrt{x''}| < \frac{1}{2\cdot 10^{m+p}},$$

so that the values of $\sqrt{x}$ and $\sqrt{x''}$, each correct to $m + p$ decimals, differ by not more than one in the last decimal place.

It has been explained that Briggs' method of computing the logarithm of a number, N, consisted in finding a root

$$\sqrt[p]{N} = 1 + x,$$

where p is of the form 2^h, and x a decimal beginning with at least fifteen ciphers, then in writing

$$\log(1 + x) = \frac{x}{a}\log(1 + a),$$

where
$$1 + a = 10^y, \qquad \text{where} \quad y = 1/2^{54}$$

so that
$$a = 12781\ 91493\ 20032\ 35 \times 10^{-32}.$$

Let us write $x = na$; since x is to begin with fifteen ciphers, n must be less than eight. Briggs' formula becomes

$$\log(1 + na) = n \log(1 + a) = \log(1 + a)^n.$$

By Taylor's Theorem, we have

$$(1 + a)^n = 1 + na + \frac{n(n-1)a^2}{2}(1 + \theta a)^{n-2}, \qquad 1 > \theta > 0,$$

whence

$$1 + na = (1 + a)^n (1 - \epsilon),$$

where
$$(1 + a)^n \epsilon = \frac{n(n-1)a^2}{2}(1 + \theta a)^{n-2}.$$

Then
$$\log(1 + na) = n \log(1 + a) + \log(1 - \epsilon).$$

Now
$$|\log(1 - \epsilon)| < \frac{M\epsilon}{1 - \epsilon},$$

and since n is less than eight, and $\dfrac{(1 + \theta a)^{n-2}}{(1 + a)^n} < 1$, we have $\epsilon < 28a^2$.

Substituting the value of a, we find $\epsilon < 46 \times 10^{-32}$, whence

$$|\log(1 - \epsilon)| < 2 \cdot 10^{-31}.$$

Then for all values of n less then eight,

$$|\log(1 + na) - n \log(1 + a)| < 2 \cdot 10^{-31}.$$

Now we may suppose that N is always between one and ten; then to find a root of the form $1 + x$, it will never be necessary to take h greater than 54.

Now is
$$\log N = p \log(1 + x) = 2^h \log(1 + x),$$

and the error in $\log N$, computed by Briggs' method, will not be greater than

$$2^{54} \times 2 \cdot 10^{-31} < \frac{1}{2 \cdot 10^{14}}.$$

Hence the inaccuracy due to the method will not prevent the result from being correct to fourteen decimals.

Note D. The Limit of Error in Weddle's Method of Calculating a Logarithm. We suppose that the number N, whose logarithm is to be calculated, lies between one and ten. We write $N = a + d$, where a is an integer less than ten, and d a decimal fraction. We divide $a + d$ by $a + 1$, unless $a + d$ begins with the figures 1.0, when the method may be advan-

tageously modified by dividing by 1.1 instead of by two. The quotient N' is then multiplied by l factors of the form $1 + c_r/10^r$, so that the product has its first k decimal figures equal to nines. Representing the product of l factors of the form $1 + c_r/10^r$ by P_l, we have

$$N'' = N' P_l = 1 - \frac{a_{k+1}}{10^{k+1}} - \frac{a_{k+2}}{10^{k+2}} \cdots .$$

It is always possible to obtain this result where $l \leqq k + 3$, and if $l = k + 3$, we shall have $a_{k+1} = 0$. The truth of this statement appears as follows: We have always $N' > .55$; in the least favorable case multiplication by four factors of the form $1 + c_r/10^r$ will raise to nines the first two figures. When the first two or more figures in the product are nines it may be necessary to multiply by two factors to raise to nine the first figure of the product different from nine, but in the least favorable case this process will raise to nine this figure and the one next following. Generally, each multiplication raises to nine one figure of the product, so that $l = k$, but we may assert that in all cases, $l \gtreqless k + 3$. An example will make these explanations clearer. Suppose

$N' = 0.735$	$c_1 = 2$
$N'P_1 = 0.8820$	$c_1' = 1$
$N'P_2 = 0.9702$	$c_2 = 2$
$N'P_3 = 0.98960\ 4$	$c_2' = 1$
$N'P_4 = 0.99950\ 004$	$c_3 = 0 \quad c_4 = 4$
$N'P_5 = 0.99989\ 98400\ 16$	$c_4' = 1$
$N'P_6 = 0.99999\ 98300\ 00$	

In this rather unfavorable case if k is two, we have to take $l = 4 = k + 2$: to make the fourth decimal a nine two factors, $1 + c_4/10^4$, are used but with the fourth not only the fifth but the sixth figures become equal to nine, so that if k is six, we have $l = 6 = k$.

Having now

$$N'' = 1 - \frac{a_{k+1}}{10^{k+1}} - \frac{a_{k+2}}{10^{k+2}} \cdots ,$$

we write

$$N'' \left(1 + \frac{a_{k+1}}{10^{k+1}}\right)\left(1 + \frac{a_{k+2}}{10^{k+2}}\right) \cdots = 1 - x,$$

and for x we find

$$x = \sum \frac{a_i a_j}{10^{i+j}} + \sum \frac{a_i a_j a_m}{10^{i+j+m}} \cdots ,$$

where $$k+1 \leqq i \leqq j \leqq m \cdots,$$

and where not more than two indices are equal in any term of the sum. Since no value of a is greater than nine, we have

$$\sum \frac{a_i a_j}{10^{i+j}} < \frac{81}{10^{2k+2}} \sum \frac{1}{10^{\alpha+\beta}} = \frac{1}{10^{2k}},$$

where α and β take independently all positive integral values. Similarly, since m is at least equal to $k+2$,

$$\sum \frac{a_i a_j a_m}{10^{i+j+m}} < \frac{729}{10^{3k+4}} \sum \frac{1}{10^{\alpha+\beta+\gamma}} = \frac{1}{10^{3k+1}}.$$

So that, finally, we have

$$x < \frac{1}{10^{2k}}\left\{1 + \frac{1}{10^{k+1}} + \frac{1}{10^{2k+2}} \cdots\right\}$$

$$< \frac{1}{10^{2k}} \frac{1}{1-\dfrac{1}{10^{k+1}}} < \frac{1.01}{10^{2k}}, \qquad \text{if } k > 1.$$

Now $$\log N'' + \sum_{r=k+1}^{\infty} \log\left(1 + \frac{a_r}{10^r}\right) = \log(1-x).$$

But $$|\log(1-x)| < \frac{Mx}{1-x} < \frac{1}{2.10^{2k}},$$

since x is known to be less than 1.01×10^{-2k}.

Then $$\log N'' = -\sum_{r=k+1}^{\infty} \log\left(1 + \frac{a_r}{10^r}\right) + \epsilon, \qquad |\epsilon| < \frac{1}{2\cdot 10^{2k}}$$

and $$\log N' = \log N'' + \sum_{l} \log\left(1 + \frac{c_r}{10^r}\right);$$

and since $\log N = \log(a+1) - \log N'$, we have

$$\log N = \log(a+1) - \sum_{l} \log\left(1 + \frac{c_r}{10^r}\right) - \sum_{r=k+1}^{\infty} \log\left(1 + \frac{a_r}{10^r}\right) + \epsilon.$$

If in the last sum we neglect all terms for which r is greater than m, we commit the error

$$\sum_{r=m+1}^{\infty} \log\left(1 + \frac{a_r}{10^r}\right) < M \sum_{r=m+1}^{\infty} \frac{a_r}{10^r} < \frac{M}{10^m} < \frac{1}{2\cdot 10^m}.$$

We compute $\log N$ from the formula

$$(4)\qquad \log N = \log'(a+1) - \sum_{l} \log'\left(1+\frac{c_r}{10^r}\right) - \sum_{r=k+1}^{m} \log'\left(1+\frac{a_r}{10^r}\right),$$

where $\log'$ means, as before, the tabulated value of a logarithm. If the tables used are correct to p places, each logarithm used is liable to an error of $1/(2\cdot 10^p)$, and we use

$$1 + l + m - k$$

tabulated values, or less if some values of a_r are zero. Now l is not greater than $k+3$, hence not more than $m+4$ logarithms are used. Then the limit of error in (4) is not greater than

$$E = \frac{1}{2}\left\{\frac{1}{10^{2k}} + \frac{1}{10^m} + \frac{m+4}{10^p}\right\}.$$

If the number of logarithms actually used is less than $m+4$, this number may replace $m+4$ in E.

We shall next consider what values may best be assigned to k and m, when p is given. It is desirable to make E as small as possible, and, in order to shorten the labor of computation, to take k no larger than is necessary. The least value possible for

$$\frac{1}{10^m} + \frac{m}{10^p}$$

is obtained by writing $m = p$. Then

$$\frac{1}{10^m} + \frac{m}{10^p} = \frac{p+1}{10^p}.$$

For clearly a value of $m \geqq p+1$ would give the expression a greater value. Whereas, if $m = p - h$,

$$\frac{1}{10^m} + \frac{m}{10^p} = \frac{10^h + p - h}{10^p} > \frac{p+1}{10^p}.$$

We shall then do best to put $m = p$, and

$$E = \frac{1}{2}\left\{\frac{1}{10^{2k}} + \frac{p+5}{10^p}\right\}.$$

To choose $2k$ less than p would make E unnecessarily large, and since E has inevitably a considerable value in the pth decimal place there is no gain in taking $2k$ greater than $p+1$. Then if p is even, let $k = p/2$, and

$$E = \frac{1}{2}\,\frac{p+6}{10^p};$$

if p is odd let $k = (p + 1)/2$, then

$$E = \frac{1}{2}\frac{p + 5.1}{10^p}.$$

In all cases if the number of logarithms actually used in (4) is less than $p + 4$, the difference may be subtracted from the numerator of E. In the example worked out in §3, we had $p = 15$, hence took $m = 15$ and $k = 8$. There were used in (4) fourteen logarithms. Hence we subtract five from the numerator of E and $E = \frac{1}{2}\frac{15.1}{10^{15}}$ and the error in the fifteenth decimal of log 5873, computed by the first method is less than 7.55. The value found, if taken correct to fourteen decimals cannot differ by more than one in the fourteenth place from the true value correct to fourteen decimals.

Now the chief part of E comes from possible inaccuracy in the tabulated logarithms. This part may be cut in half by a slight modification suggested and used by Gernerth. With each logarithm tabulated it is indicated whether the given value is smaller or larger than the true value. If the tabulated value, l', is too small, the true value, l, is between l' and $l' + 1/(2\cdot10^{15})$, and if in place of l' we use $l' + 1/(4\cdot10^{15})$, the difference from l cannot be more than $1/(4\cdot10^{15})$. Similarly, if the tabulated value is too large, and in place of l' we use $l' - 1/(4\cdot10^{15})$, the difference from l is not more than $1/(4\cdot10^{15})$. If this modification is practised, we have

$$E' = \frac{1}{4}\left\{\frac{2}{10^{2k}} + \frac{2}{10^m} + \frac{m + 4}{10^p}\right\},$$

E' being now the limit of error in (4).

It appears that m should be given the value $p + 1$. The last logarithm in the sum, $\log'\left(1 + \frac{a_{p+1}}{10^{p+1}}\right)$, will be tabulated as zero, for

$$\log\left(1 + \frac{a_{p+1}}{10^{p+1}}\right) < \log\left(1 + \frac{1}{10^p}\right) < \frac{M}{10^p} < \frac{1}{2\cdot10^p},$$

but in the modified calculation, its value must be written as $1/(4\cdot10^p)$. As before, k should be chosen as $(p + 1)/2$ or $p/2$, according as p is odd or even. If p is even,

$$E' = \frac{1}{4}\frac{p + 7.2}{10^p};$$

$$\text{if } p \text{ is odd, } E' = \frac{1}{4}\frac{p + 5.4}{10^p}.$$

If the number of logarithms actually used is less than $p + 5$, the difference may be subtracted from the numerator of E'. In the example of §3, using this second method, fifteen logarithms are used, so that $E' = \frac{3.85}{10^{15}}$. But even this result does not enable us to say that the value of log 5873, computed by the second method, and taken correct to fourteen decimals is equal to the true value correct to fourteen decimals.

It remains to consider one more point. In order to determine the values of $a_k, a_{k+1}, \cdots a_p$, to how many places must the division of N by $a + 1$, and the multiplication of the quotient by the factors $1 + c_r/10^r$ be carried?

Suppose all this work to be carried out correctly to n decimals. Then the error in the value used for N' is not greater that $1/(2\cdot 10^n)$. This error is multiplied by $c_r/10^r$ and the product taken correct to n decimals, so that, in the value used for $N'(c_r/10^r)$, the limit of error is $\frac{1}{2\cdot 10^n}\left(1 + \frac{c_r}{10^r}\right)$. Then in the value found to n places for $N'\left(1 + \frac{c_r}{10^r}\right)$ the limit of error is

$$\frac{1}{2\cdot 10^n}\left\{1 + \left(1 + \frac{c_r}{10^r}\right)\right\}.$$

Repetition of the reasoning shows us that the limit of error in $N'P_l$ is

$$\frac{1}{2\cdot 10^n}\left\{1 + \left(1 + \frac{c_s}{10^s}\right) + \cdots + P_l\right\} < \frac{(l+1)P_l}{2\cdot 10^n},$$

where $1 + \frac{c_s}{10^s}$ is the last factor in P_l. We have alway P_l less than two, since $N' > .55$ and $N'P_l < 1$; generally P_l is very much less than two, and since $l < k + 3 < \frac{p+1}{2} + 3$, the error in $N'P_l$ will be less than $\frac{1}{2}\frac{p+9}{10^n}$. It will then suffice generally to choose $n = p + 2$ to have a correct value of a_p. In the example of §3, l is 7, and $P_l < \frac{10}{9}$, so that the error is less than

$$\frac{8.\frac{10}{9}}{2\cdot 10^n} < \frac{4.5}{10^n}.$$

To have the fifteenth place correct it is here sufficient to take $n = 17$.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS., APRIL, 1907.

CUBIC CURVES IN RECIPROCAL TRIANGULAR SITUATION

By John Frederick Messick

Introduction. The purpose of this paper is to find two cubic curves in what we shall call reciprocal triangular situation; that is, two cubics so related to each other that a single infinity of triangles with vertices on one may have their sides tangent to, or as we say, on the other, and also that the sides of a singly infinite system of triangles with vertices on the second may be on the first.

1. The Curve of Order and Class Three. We shall start with a cuspidal cubic and write it in the form

$$x_1x_3^2 = x_2^3$$

to which any cuspidal cubic may be reduced.*

We see at once from this equation that, taking the triangle with sides x_1, x_2, x_3, and vertices 1, 2, 3 as reference triangle, there is a cusp at the vertex 1, a flex at 3, x_3 is the cusp tangent, and x_1 the flex tangent. But we wish to see this from the parametric form of the cubic. This is

$$x_1 : x_2 : x_3 = \lambda^3 : \lambda : 1,$$

since the coördinates of any point on the cubic may be expressed as λ^3, λ, 1, where λ is a variable parameter.

Substituting these in the equation of any line $(\xi x) \equiv \xi_1x_1 + \xi_2x_2 + \xi_3x_3 = 0$, we get

$$\xi_1\lambda^3 + \xi_2\lambda + \xi_3 = 0.$$

The λ^2 term is lacking; therefore the sum of the roots of this equation, say λ_1, λ_2, λ_3, vanishes; and the parameters of the three points on the cubic in which it is cut by a straight line are connected, therefore, by the relation

$$\lambda_1 + \lambda_2 + \lambda_3 = 0;$$

or as we shall say, this is the involution for points of the cubic lying on a straight line.

* Salmon, *Higher Plane Curves*, page 180.

The triple points of the involution are given when the λ's become equal, or by

$$3\lambda = 0,$$

the three roots of which are ∞, ∞, 0.

Assuming only that the cubic is rational, we know that the possible cases in which a line may cut a cubic in three coincident points, that is, at triple points, are the two nodal tangents, the cusp tangent, and the flex tangent. The possibilities then for singular points are node, cusp, and flex. Any line through a node or cusp cuts there in two coincident points. In the case of a node, when the curve is expressed parametrically, these two coincident points are given by distinct values of the parameter. But from the roots of our equation the two values of the parameter are the same, namely ∞. We infer, therefore, that if the equation giving the triple points has two equal roots, the cubic has a cusp given by this double value of the parameter and a flex given by the other root. Thus, as is obvious from the equation for a rational cubic, the triple points of the involution give the flexes, the neutral pair give the node. For a cuspidal cubic the neutral pair are equal, and two triple points are also equal. Thus, the condition for a cuspidal cubic is that the equation for the triple points of the involution shall have equal roots. Our cubic, therefore, has a cusp for the double root ∞, that is, at the point 1, 0, 0, and a flex for the root 0 at 0, 0, 1.

From these considerations we shall be able to determine readily the character of a cubic appearing in a more complicated form.

2. The Cubics ϕ and f. If a conic be passed through three given points on the cubic, it will cut the cubic in three other points. The six points in which the conic cuts the cubic are obtained by substituting the parameters λ^3, λ, 1 for x_1, x_2, x_3 respectively in the equation of the conic $(\xi x)^2 = 0$. This gives

$$\xi_1^2\lambda^6 + 2\xi_1\xi_2\lambda^4 + 2\xi_1\xi_3\lambda^3 + \xi_2^2\lambda^2 + 2\xi_2\xi_3\lambda + \xi_3^2 = 0.$$

Here the term in λ^5 is wanting; the sum of the roots of the equation, say $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6$, is therefore equal to zero. We have then as a symmetrical relation of the λ's independent of the ξ's

$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 = 0,$$

which we may say is the involution along the cubic for points on a conic.

Since we suppose three of the points fixed, the sum of their parameters therefore being a constant, we may say that the involution for the other three is

$$\lambda_1 + \lambda_2 + \lambda_3 = \kappa,$$

which we shall call an I_1^3, meaning by this symbol an involution of three things in which two must be given to determine the third. By I_2^3 is meant an involution in which, if one is given, the other two are determined.

The conic expressed in terms of symmetric functions of the parameters of points which it cuts out is

$$x_1^2 + S_2x_1x_2 - S_3x_1x_3 + S_4x_2^2 - S_5x_2x_3 + S_6x_3^2 = 0.$$

By breaking up the six points into two sets of three, say σ and σ', the involution already found is

$$\sigma_1 + \sigma_1' = 0\,;$$

or substituting $-\kappa$ for one of the sets which we suppose fixed, σ' say, we have it in the form given,

$$\sigma_1 = \lambda_1 + \lambda_2 + \lambda_3 = \kappa.$$

Now by putting the conic through a point a not on the cubic, we get $a_1^2 + a_1a_2(\sigma_2 + \sigma_1\sigma_1' + \sigma_2') - a_1a_2(\sigma_3 + \sigma_2\sigma_1' + \sigma_1\sigma_2' + \sigma_3') + a_2^2(\sigma_3\sigma_1' + \sigma_2\sigma_2' + \sigma_1\sigma_3') - a_2a_3(\sigma_3\sigma_2' + \sigma_2\sigma_3') + a_3^2\sigma_3\sigma_3' = 0$. Since σ' is fixed this may be put in the form

$$A_0\sigma_3 + A_1\sigma_2 + A_2\sigma_1 + A_3 = 0,$$

where the A's are functions of the known quantities a and σ'. And since σ' is a constant, this may be written, if we exclude the special case when $A_0 = 0$,* as

$$\lambda_1\lambda_2\lambda_3 - a(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1) - a\beta\kappa = 0,$$

a second relation among the varying λ's, also an I_1^3.

Through four given points — three on and one off the cubic — a whole pencil of conics may be passed, each cutting the cubic in three other points; hence, a single infinity of triads, or sets of three points, arise. If these be joined, a system of lines is given, the envelope of which is the curve we wish to obtain.

* This means that a value of the parameter is ∞; that is, the point is taken at the cusp, which case we shall except.

This system of triangles is such that, if one side is given, the other two are determined. If we take a line, therefore, through two points of the cubic, and subject it to the two involutions

$$\lambda_1 + \lambda_2 + \lambda_3 = \kappa,$$

$$\lambda_1\lambda_2\lambda_3 - a(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1) - a\beta\kappa = 0,$$

the combination of which is an I_2^3, the locus of the line will be the required curve.

The points of the varying triad, as we have supposed, are given by the parameters λ_1, λ_2, λ_3; and the equation of a line through two of these, say those given by λ_1 and λ_2, is, writing it in determinant form,

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ \lambda_1^3 & \lambda_1 & 1 \\ \lambda_2^3 & \lambda_2 & 1 \end{vmatrix} = 0,$$

which becomes on reduction

$$x_1 - (\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2)x_2 + \lambda_1\lambda_2(\lambda_1 + \lambda_2)x_3 = 0.$$

And now if we put the conditions given by the two involutions (the two I_1^3's) on this line, as the conic varies the line will describe the curve in question.

Put λ for λ_3 in the involutions, and eliminate λ_1 and λ_2 from the three equations. The result is

$$(\lambda - a)x_1 - \{\lambda(\lambda-\kappa)^2 - a\beta\kappa + a\kappa\lambda - a\kappa^2\}x_2$$
$$+ a\{\lambda(\lambda-\kappa)^2 - \beta\kappa(\lambda-\kappa)\}x_3 = 0.$$

This for a fixed λ is a line of the required curve; and for varying λ it gives the curve itself. And since it is of the third degree in λ, it is a cubic. Also, since three lines may be drawn from a point tangent to the curve, it is of the third class.

We have now the two cubics, the original in points,

$$\phi: \begin{cases} x_1 = \lambda^3 \\ x_2 = \lambda \\ x_3 = 1, \end{cases}$$

and the other, also written in parametric form, in lines,

$$f:\quad \begin{cases} \xi_1 = \lambda - a \\ \xi_2 = -\{\lambda(\lambda-\kappa)^2 + a\kappa(\lambda-\beta-\kappa)\} \\ \xi_3 = a\{\lambda(\lambda-\kappa)^2 - \beta\kappa(\lambda-\kappa)\}, \end{cases}$$

so related that triangles may have their vertices on ϕ and their sides on f; or as we shall say, they lie in triangular situation.

To get the relation of the parameters of three points on ϕ, such as λ_3, the parameter of which gives the line through λ_1 and λ_2, when the lines of f given by these parameters meet in a point, substitute the ξ's of f in the equation of a point $(x\xi) = 0$. This gives, writing it in powers of λ,

$$(x_2 - ax_3)\lambda^3 - 2\kappa(x_2 - ax_3)\lambda^2 - (x_1 - \kappa^2x_2 - a\kappa x_2 + a\kappa^2x_3 - a\beta\kappa x_3)\lambda$$
$$+ ax_1 - a\beta\kappa x_2 - a\kappa^2x_2 - a\beta\kappa^2x_3 = 0.$$

Comparing this with the equation of symmetric functions, we get as a symmetrical relation of the λ's independent of the x's

$$\lambda_1 + \lambda_2 + \lambda_3 = 2\kappa.$$

The triple lines of this involution are given by

$$3\lambda = 2\kappa,$$

the three roots of which are ∞, ∞, $\frac{2\kappa}{3}$. Here again, since the cubic is rational, and the equation giving the triple lines has a double root, it is a curve of class three with a flex, that is, a cuspidal cubic. As this is the dual of the other case, here the double root ∞ gives the flex tangent, and the single root $\frac{2\kappa}{3}$ gives the cusp tangent of f.

On the substitution of ∞ in the equation of ϕ the flex tangent is seen to be

$$x_2 - ax_3 = 0.$$

Since x_1 is missing, this is a line through the reference point 1, 0, 0. The flex line of f, therefore, passes through the cusp of ϕ.

3. The Cubics ϕ and f'. We wish now to put the condition on f such that the triangular relation between it and ϕ will be reciprocal. To do this we shall take ϕ in lines and get another cubic so related to it that tri-

angles will have vertices on it and sides on ϕ; and then identify this new cubic, which we shall call f', with f.

We get the line equation of ϕ by taking a line through two of its points λ_1 and λ_2 and letting the points come together at λ_1, say; that is, by making the λ's equal in the equation

$$x_1 - (\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2)x_2 + \lambda_1\lambda_2(\lambda_1 + \lambda_2)x_3 = 0.$$

This gives the tangent at the point λ_1,

$$x_1 - 3\lambda_1^2 x_2 + 2\lambda^3 x_3 = 0,$$

which for varying λ is the line equation of ϕ itself.

We may also get this by the determinant scheme which we shall make use of in what follows, namely,

$$\xi_i = \begin{vmatrix} x_j & x_k \\ x_j' & x_k' \end{vmatrix} \qquad (i, j, k = 1, 2, 3,\ 2, 3, 1,\ 3, 1, 2)$$

where x' is the derivative of x with respect to λ. This holds as well for changing from lines to points as from points to lines by simply interchanging ξ's and x's. This scheme gives, as before, for the parametric form of ϕ in lines

$$\xi_1 : \xi_2 : \xi_3 = 1 : -3\lambda^2 : 2\lambda^3.$$

To get the relation of the parameters of the three points the tangents at which pass through a point we substitute these values in the equation of a point. This gives

$$x_1 - 3\lambda^2 x_2 + 2\lambda^3 x_3 = 0,$$

in which the term in λ is missing. This tells us that the sum of the products of the roots by two's is zero, that is,

$$\lambda_2\lambda_3 + \lambda_3\lambda_1 + \lambda_1\lambda_2 = 0,$$

or

$$\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} = 0.$$

The triple lines of this involution are given by the roots of

$$\frac{3}{\lambda} = 0,$$

from which we have

$$\lambda = 0,\ 0,\ \infty.$$

The double root 0 gives the flex tangent, that is, the line $x_1 = 0$; and the single root ∞ gives the cusp tangent $x_3 = 0$.

Substituting the parametric values of ξ in the conic $(x\xi)^2 = 0$, and arranging in powers of λ, we get

$$4x_3^2\lambda^6 - 12x_2x_3\lambda^5 + 9x_2^2\lambda^4 + 4x_1x_2\lambda^3 - 6x_1x_2\lambda^2 + x_1^2 = 0,$$

a sextic in λ which gives the six common tangents to the cubic and conic; that is to say, the roots of this equation are the parameters of the points on the cubic ϕ at which these common lines are tangent to the cubic. This equation has no first power term in λ, therefore the sum of the products of its roots by fives vanishes; or

$$\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} + \frac{1}{\lambda_4} + \frac{1}{\lambda_5} + \frac{1}{\lambda_6} = 0.$$

Fixing three of these, say $\lambda_4, \lambda_5, \lambda_6$, we have as an involution for the other three

$$\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} = \kappa_1;$$

that is, if we take tangents at three fixed points of ϕ, and take a conic tangent to these, the points of tangency of the other three common lines are connected by this relation.

Taking three lines on ϕ and one off, analogously to the other case where we had ϕ in points and took three points on and one off, also making use of the involution above, we get as a second relation among the λ's for this case

$$\frac{1}{\lambda_1\lambda_2\lambda_3} - a_1\left(\frac{1}{\lambda_1\lambda_2} + \frac{1}{\lambda_2\lambda_3} + \frac{1}{\lambda_3\lambda_1}\right) - a_1\beta_1\kappa_1 = 0.$$

By subjecting the point of intersection of two lines of ϕ to these two involutions the combination of which is likewise an I_2^3, we shall get a locus such that triangles with vertices on it will have their sides on ϕ; since it is a locus of triads related to ϕ in such a way that if one point of the triad is given the other two are determined.

The intersection of two lines of ϕ is

$$\begin{vmatrix} \xi_1 & \xi_2 & \xi_3 \\ 1 & -3\lambda_1^2 & \lambda_1^3 \\ 1 & -3\lambda_2^2 & \lambda_2^3 \end{vmatrix} = 0,$$

which reduces to

$$6\lambda_1^2\lambda_2^2\xi_1 + 2(\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2)\xi_2 + 3(\lambda_1 + \lambda_2)\xi_3 = 0.$$

Subject this point to the conditions expressed in the two involutions; that is, eliminate λ_1 and λ_2 from the three equations and we get, putting λ for λ_3,

$$\begin{aligned}6(\lambda^2 - a_1\lambda^3)\xi_1 - 2\{(a_1\kappa_1^2 + a_1\beta_1\kappa_1)\lambda^3 - (a_1\kappa_1 + \kappa_1^2)\lambda^2 + 2\kappa_1\lambda - 1\}\xi_2 \\ + 3\{a_1\beta_1\kappa_1^2\lambda^3 + (a_1\kappa_1^2 - a_1\beta_1\kappa_1)\lambda^2 - 2a_1\kappa_1\lambda + a_1\}\xi_3 = 0,\end{aligned}$$

which for fixed λ is a point on the curve f', for varying λ the curve itself. Writing it in powers of $1/\lambda$,

$$\begin{aligned}(2\xi_2 + 3a_1\xi_3)\frac{1}{\lambda^3} - 2\kappa_1(2\xi_2 + 3a_1\xi_3)\frac{1}{\lambda^2} + \{6\xi_1 + 2(a_1\kappa_1 + \kappa_1^2)\xi_2 \\ + 3(a_1\kappa_1^2 - a_1\beta_1\kappa_1)\}\frac{1}{\lambda} - \{6a_1\xi_1 + 2(a_1\kappa_1^2 + a_1\beta_1\kappa_1)\xi_2 - 3a_1\beta_1\kappa_1^2\}\xi_3 = 0,\end{aligned}$$

from which we get the symmetrical relation of the λ's

$$\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} = 2\kappa_1.$$

The condition for triple points in this involution is

$$\frac{3}{\lambda} = 2\kappa_1.$$

The solution of this gives

$$\lambda = 0,\ 0,\ \frac{3}{2\kappa_1},$$

which tells us that this is likewise a cuspidal cubic, the cusp being given for the double root 0, and the flex for the other value of the root $\frac{3}{2\kappa_1}$.

This cubic in parametric form is

$$f': \begin{cases} x_1 = 6(\lambda^2 - a_1\lambda^3) \\ x_2 = -2\{a_1\kappa_1^2 + a_1\beta_1\kappa_1)\lambda^3 - (a_1\kappa_1 + \kappa_1^2)\lambda^2 + 2\kappa_1\lambda - 1\} \\ x_3 = 3\{a_1\beta_1\kappa_1^2\lambda^3 + (a_1\kappa_1^2 - a_1\beta_1\kappa_1)\lambda^2 - 2a_1\kappa_1\lambda + a_1\}. \end{cases}$$

4. Identification of f and f'. We wish now to identify f' with f: for, since f is so related to ϕ that triangles with vertices on ϕ have their sides tangent to f, and f' is so related to ϕ that the sides of triangles with vertices

on it are tangent to ϕ, by bringing these two, f and f', together we shall have a cubic which has the reciprocal relation with ϕ; or as we say, the two stand in the reciprocal triangular relation or situation.

We have seen that the flex tangent of f goes through the cusp of ϕ. We must put the condition on f, therefore, that will make its cusp lie on the flex tangent of ϕ. This will give us a relation between a, β, and κ.

To get this relation we change f from lines to points; and then, putting in the condition for cusp, namely, $\lambda = \frac{2\kappa}{3}$, we equate the first coordinate x_1 to zero. This gives

$$x_1 \equiv \begin{vmatrix} -2\kappa^3 + 9a\kappa^2 + 27a\beta\kappa, & 2\kappa^3 + 9\beta\kappa^2 \\ \kappa^2 - 3a\kappa, & -\kappa^2 - 3\beta\kappa \end{vmatrix} = 0,$$

which reduces to

$$(a + \beta)\kappa^2 + 9a\beta\kappa + 27a\beta^2 = 0. \tag{1}$$

As we shall see, f and f' will each have three contacts with ϕ. If these contacts be identified, and also the flex points of the two, the cubics themselves will be completely identified.

We shall first look at the condition for the intersections of the two cubics ϕ and f. Since the points of the triad λ_1, λ_2, λ_3 lie on ϕ and the sides lie on

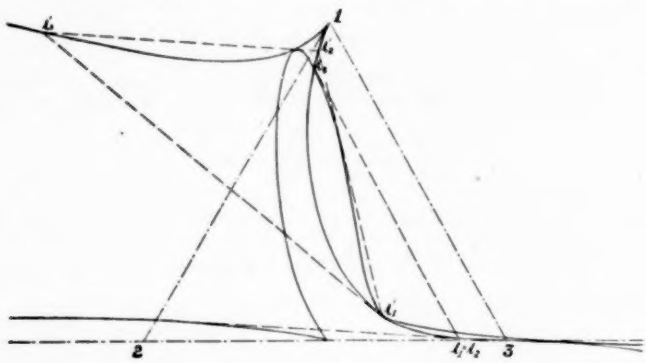


Fig. 1.

f, if we let two of the points, say λ_1 and λ_2, come together, thereby making the lines $\lambda_1\lambda_3$ and $\lambda_2\lambda_3$, which always intersect on ϕ and are tangent to f, coincide, in the limit the points of tangency of the lines with f and the intersection of the lines coincide on ϕ, of course; and the double line becomes tangent to f

at its point of intersection with ϕ. The line $\lambda_1\lambda_2$ has become a tangent to ϕ where this double line cuts ϕ. The involution then becomes

$$2\lambda_1 + \lambda_3 = \kappa ;$$

that is, this is the condition for intersections of ϕ and f.

In the case of contacts one of the sides of the triangle, $\lambda_1\lambda_3$ say, becomes a common tangent to the two curves at a vertex, say λ_1.* The two curves have, as we say, a common tangent $\lambda_1\lambda_3$ at this point.† We have, therefore, as the condition for contacts between ϕ and f

$$2\lambda_1 + \lambda_3 = 0.$$

Eliminating λ_2 and λ_3 from the three equations

$$2\lambda_1 + \lambda_3 = 0,$$
$$\lambda_1 + \lambda_2 + \lambda_3 = \kappa,$$
$$\lambda_1\lambda_2\lambda_3 - a(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1) - a\beta\kappa = 0,$$

putting λ for λ_1, we get for the equation giving the contacts of ϕ and f

$$(2) \qquad 2\lambda^3 + (2\kappa - 3a)\lambda^2 - a\kappa\lambda + a\beta\kappa = 0,$$

a cubic equation in λ, showing, therefore, that there are three points of contact between the two cubics.

Likewise the contacts of ϕ and f' are given by the condition

$$\frac{2}{\lambda_1} + \frac{1}{\lambda_3} = 0.$$

Eliminating λ_2 and λ_3 from this and the equations of the two involutions

$$\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} = \kappa_1,$$

$$\frac{1}{\lambda_1\lambda_2\lambda_3} - a_1\left(\frac{1}{\lambda_1\lambda_2} + \frac{1}{\lambda_2\lambda_3} + \frac{1}{\lambda_3\lambda_1}\right) - a_1\beta_1\kappa_1 = 0,$$

putting as before λ for λ_1, we get

$$(3) \qquad a_1\beta_1\kappa_1\lambda^3 - a_1\kappa_1\lambda^2 + (2\kappa_1 - 3a_1)\lambda + 2 = 0,$$

the equation giving the three contacts of ϕ and f'.

* See triangle $\lambda_1'\lambda_2'\lambda_3'$ in figure 1.

† For proof of these for the general case see F. Morley: On Two Cubic Curves in Triangular Relation, *Proceedings of the London Mathematical Society*, ser. 2, vol. 4, part 5.

Comparing equations (2) and (3), we see that in order for these to be the same, that is, in order for f and f' to have the same contacts with ϕ, we have the relation among the coefficients

$$\frac{2}{a_1\beta_1\kappa_1} = -\frac{2\kappa - 3a}{a_1\kappa_1} = -\frac{a\kappa}{2\kappa_1 - 3a_1} = \frac{a\beta\kappa}{2}, \tag{4}$$

from which we infer that the constants with subscripts may be interchanged with those without them.

We shall next identify the flex points of f and f'. We get the flex point of f by writing it in points and putting in the condition for a flex, namely, $\lambda = \infty$. By the determinant scheme the coordinates of any point x on f are seen to be

$$x_1 = -a \begin{vmatrix} \lambda(\lambda-\kappa)^2 + a\kappa\,(\lambda - \beta - \kappa), & \lambda(\lambda-\kappa)^2 - \beta\kappa(\lambda-\kappa) \\ (\lambda-\kappa)^2 + 2\lambda(\lambda-\kappa) + a\kappa, & (\lambda-\kappa)^2 + 2\lambda(\lambda-\kappa) - \beta\kappa \end{vmatrix}$$

$$x_2 = a \begin{vmatrix} \lambda(\lambda-\kappa)^2 - \beta\kappa(\lambda-\kappa), & \lambda - a \\ (\lambda-\kappa)^2 + 2\lambda(\lambda-\kappa) - \beta\kappa, & 1 \end{vmatrix}$$

$$x_3 = -\begin{vmatrix} \lambda - a, & \lambda(\lambda-\kappa)^2 + a\kappa(\lambda-\beta-\kappa) \\ 1, & (\lambda-\kappa)^2 + 2\lambda(\lambda-\kappa) + a\kappa \end{vmatrix}.$$

Putting $\lambda = \infty$ in these, we find that they reduce, respectively, to

$$a\kappa(a+\beta),\ a,\ 1,$$

which are the coordinates of the flex point of f. Or we have as the equation giving this flex point

$$a\kappa(a+\beta)\xi_1 + a\xi_2 + \xi_3 = 0. \tag{5}$$

The flex point of f' is obtained by putting $\lambda = \dfrac{3}{2\kappa_1}$ in the equation of f'. When this value of λ is put in, the equation of f' reduces to

$$54(2\kappa_1 - 3a_1)\xi_1 - 2\kappa_1(9a_1\kappa_1 - 2\kappa_1^2 + 27a_1\beta_1)\xi_2 + 3\kappa_1(9a_1\beta_1\kappa_1 + 2a_1\kappa_1^2)\xi_3 = 0. \tag{6}$$

If now these two flex points are to be the same, we must have the relation

$$\frac{54(2\kappa_1 - 3a_1)}{a\kappa(a+\beta)} = -\frac{2\kappa_1(9a_1\kappa_1 - 2\kappa_1^2 + 27a_1\beta_1)}{a} = \frac{3\kappa_1(9a_1\beta_1\kappa_1 + 2a_1\kappa_1^2)}{1}. \tag{7}$$

Hence, this is the condition that the flexes of f and f' coincide.

The conditions expressed in (4) and (7) are sufficient to identify the two cubics completely; for, to say they shall have contact at the same point is two conditions, therefore six conditions for the three contacts; and identifying the flex points is one other condition, making in all seven. But seven conditions completely fix a cuspidal cubic, since requiring it to have a cusp is equivalent to two conditions.

5. Equation of f in terms of β. We have already found one relation connecting a, β, and κ, namely, equation (1). From (4) and (7) we shall find another, and by means of these two shall be able to express two of the constants in terms of the third.

To avoid writing subscripts interchange the constants a_1, β_1, κ_1 with a, β, κ, which we saw by (4) could be done, in expressions (4) and (7), and eliminate a_1, β_1, κ_1 from the resulting expressions, namely,

$$(4') \qquad \frac{2}{a\beta\kappa} = -\frac{2\kappa_1 - 3a_1}{a\kappa} = -\frac{a_1\kappa_1}{2\kappa - 3a} = \frac{a_1\beta_1\kappa_1}{2},$$

$$(7') \qquad \frac{54(2\kappa - 3a)}{a_1\kappa_1(a_1 + \beta_1)} = -\frac{2\kappa(9a\kappa - 2\kappa^2 + 27a\beta)}{a_1} = 3\kappa^2(9a\beta\kappa + 2a\kappa^2).$$

This elimination gives

$$(8) \qquad 8\kappa^3 - 60a\kappa^2 + 54a^2\kappa - 108a\beta\kappa + 81a^2\beta = 0.$$

And now eliminating a from (1) and (8), we get

$$(9) \qquad 8\kappa^4 + 204\beta\kappa^3 + 1782\beta^2\kappa^2 + 6561\beta^3\kappa + 8748\beta^4 = 0,$$

a homogeneous equation of the fourth degree in β and κ.

The solution of this equation gives $\dfrac{\kappa}{\beta} = -\dfrac{9}{2}$ or -12, and by putting these in (1) we find also $\dfrac{a}{\beta} = -3$ or $-\dfrac{16}{7}$. Taking the last set of these, namely

$$a = -\frac{16}{7}\beta, \quad \kappa = -12\beta,$$

and substituting in the original equation of f, we get it in the form

$$f: \begin{cases} \xi_1 = \lambda + \dfrac{16}{7}\beta \\ \xi_2 = -\lambda^3 - 24\beta\lambda^2 - \dfrac{1200}{7}\beta^2\lambda - \dfrac{2112}{7}\beta^3 \\ \xi_3 = -\dfrac{16}{7}\beta\lambda^3 - \dfrac{384}{7}\beta\lambda^2 - \dfrac{2496}{7}\beta^3\lambda - \dfrac{2304}{7}\beta^4. \end{cases}$$

6. Constants and Conditions. We have found that a and κ may be expressed in terms of β, which may be given an infinite number of values; and each of these gives a cubic in the desired relation. There is therefore a single infinity of cubics satisfying the requirements, namely, having this reciprocal triangular relation with the original cubic.

This follows also from consideration of the conditions imposed. To say a cubic shall have a cusp imposes two conditions upon it, leaving seven conditions, or seven degrees of freedom for the cuspidal cubic. To say there shall be three contacts is equivalent to three conditions, and to make this reciprocal puts on three more. There is left then one degree of freedom, which says likewise that there is a single infinity of cubics in the proper situation.

7. Contacts of ϕ and f. The equation giving the contacts of ϕ and f we have obtained as

$$2\lambda^3 + (2\kappa - 3a)\lambda^2 - a\kappa\lambda + a\beta\kappa = 0.$$

On substitution for κ and a in terms of β this becomes

$$7\lambda^3 - 60\beta\lambda^2 - 96\beta^2\lambda + 96\beta^3 = 0.$$

The discriminant of this equation is of the form

$$D = A\beta^6,$$

where A is a positive number, and is positive for all values of β. The equation, therefore, has three real roots whatever value we give to β; that is, the curves always have three real contacts.

For β positive there are two positive roots and one negative, that is, two contacts with the part of ϕ which lies within the triangle of reference, and one for the part outside. For β negative there is one contact within and two without the triangle; since one root is positive and two are negative.

We shall see that $\beta = \pm n$ gives the cusps of f on x_1 harmonic with the reference points 0, 1, 0 and 0, 0, 1. So here, the contacts given by $\beta = \pm n$ lie on a line through 0, 0, 1 and are harmonic with this point and the point where x_3 cuts the line; that is to say, lines through 0, 0, 1 and the points of contact for $\beta = n$ cut ϕ again at the points of contact for $\beta = -n$.

8. Intersections of ϕ and f. To get the equation giving the intersections we eliminate λ_1 and λ_2 from

$$2\lambda_1^2 + \lambda_3 = \kappa,$$

which is the condition for intersections, where λ_3 is the parameter of this point, and the involutions

$$\lambda_1 + \lambda_2 + \lambda_3 = \kappa,$$

$$\lambda_1\lambda_2\lambda_3 - a(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1) - a\beta\kappa = 0.$$

Calling λ_3, λ, this gives

$$\lambda^3 + (3a - 2\kappa)\lambda^2 + (\kappa^2 - 2a\kappa)\lambda - a\kappa^2 - 4a\beta\kappa = 0.$$

The cubics, therefore, have three intersections.

Making the substitution $a = -\frac{16}{7}\beta$, $\kappa = -12\beta$ in this equation, we get

$$7\lambda^3 + 120\beta\lambda^2 + 624\beta^2\lambda + 1536\beta^3 = 0.$$

The discriminant of this is of the form

$$D = -A'\beta^6:$$

that is, it is negative whatever value β may have, which tells us there is only one real root. The cubics, therefore, have but one real intersection.

9. Characteristics of the System. As different values are given to β in the equation of f, the cusp of f takes different positions along the line x_1. As β varies from 0 to $-\infty$ the cusp passes from the point 0, 0, 1 to 0, 1, 0, the cusp tangent passing from coincidence with x_1 to coincidence with x_3, and the flex tangent at the same time passing from coincidence with x_2 to coincidence with x_3. While β is taking all these negative values, there is one contact in the triangle and two outside; and the real intersection is in the triangle.

For positive values of β the cusp traverses the external segment of the side x_1, passing from the point 0, 1, 0 around to 0, 0, 1 through infinity as β diminishes from $+\infty$ to 0. Here there are two contacts in the triangle and one outside. Also, in this case the intersection of ϕ and f is outside.

The cusp of f is at the middle point of the base for $\beta = -\frac{1}{4}$. For $\beta = \frac{1}{4}$ it is at infinity, the cusp tangent being parallel to the base for this value. In general $\beta = \pm n$ gives two positions of the cusp, that is, two points on x_1 which are harmonic with 0, 1, 0 and 0, 0, 1.

10. Degenerate cases of f. For certain values of β the cubic f will degenerate, and we wish to consider these degenerate cases.

When the cusp tangent is x_1 and the flex tangent is x_2; that is, when $\beta = 0$, the cusp and flex coincide at their intersection 0, 0, 1. The cubic,

therefore, must break up for this value of β. To show how it breaks up substitute $d\beta$ for β in the equation of f. Disregarding higher powers of $d\beta$ than the first, we write f simply

$$\xi_1 = \lambda + \frac{16}{7}d\beta, \quad \xi_2 = -\lambda^3 - 24d\beta\lambda^2, \quad \xi_3 = -\frac{16}{7}d\beta\lambda^3;$$

or let us say

$$\xi_1 = \lambda + a, \quad \xi_2 = -\lambda^3 - b\lambda^2, \quad \xi_3 = -c\lambda^3,$$

where a, b, c are multiples of $d\beta$. Building up a vanishing function in these infinitesimals, we get the equation

$$ac^2\xi_2^3 - \frac{a+b}{c}\xi_3^3 - b^3\xi_1\xi_2^2 - (3ac + bc)\xi_2^2\xi_3 + (3a + 2b)\xi_2\xi_3^2 = 0.$$

Now making $d\beta$ vanish, since the coefficient of ξ_3^3 is the only one independent of $d\beta$, $d\beta$ dividing out and leaving only a numerical quantity, we have simply

$$\xi_3^3 = 0,$$

which is the point 0, 0, 1 three times over.

Likewise, when $\beta = \infty$, the cusp tangent coincides with the flex tangent in x_3. The cusp falls at 0, 1, 0 and the flex at 1, 0, 0; and the cubic again breaks up.

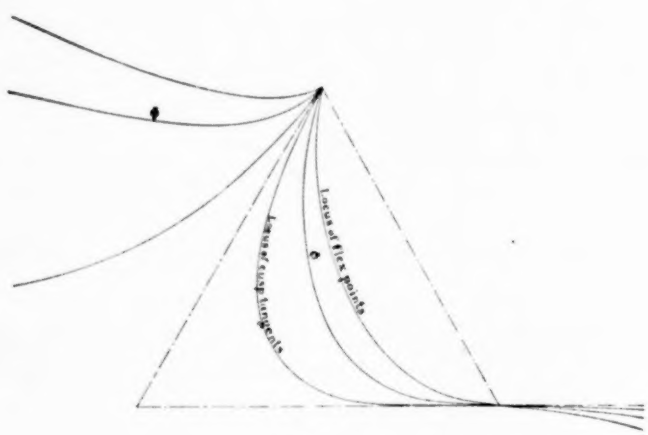


FIG. 2.

11. Locus of the Cusp Tangent and Flex. If the condition for cusp tangent, $\lambda = \frac{2}{3}\kappa = -8\beta$, is put in f, this reduces to the cusp tangent,

and gives as its equation

$$\xi_1 = -1, \quad \xi_2 = 8\beta^2, \quad \xi_3 = 32\beta^3.$$

This for a given β is the cusp tangent of the particular cubic of the system given by that β. For varying β it is the locus of the cusp tangents of the whole system.

From its form we recognize this as a cubic which has a cusp for the double root $\beta = 0$ and a flex for $\beta = \infty$; that is, it has its cusp and flex at 1, 0, 0 and 0, 0, 1, respectively, the same as the original cubic.

The flex point of f we have found was

$$x_1 = a\kappa(a + \beta), \quad x_2 = a, \quad x_3 = 1,$$

which in terms of β is

$$x_1 = -1728\beta^3, \quad x_2 = -112\beta, \quad x_3 = 49.$$

Its locus for varying β is likewise another cubic with cusp and flex coincident with those of ϕ.

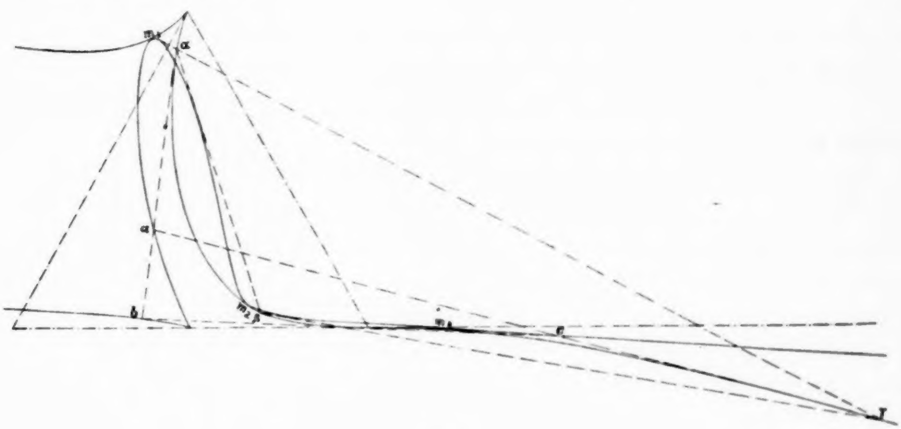


FIG. 3.

12. The Figure. In figure 3 the cubic f is drawn for the case $\beta = -\frac{1}{4}$; its cusp, therefore, lies at the middle point of x_1. The cusp tangent

$$x_1 - 8\beta x_2 - 32\beta^3 x_3 = 0$$

becomes

$$2x_1 - x_2 + x_3 = 0$$

and cuts x_3, therefore, at one third the distance from 0, 1, 0 to 1, 0, 0.

The flex tangent

$$x_2 + \frac{16}{7}\beta x_3 = 0$$

becomes

$$7x_2 - 4x_3 = 0$$

and passes through 1, 0, 0 and cuts x_1 in the ratio 7 : 4. The flex point

$$1728\beta^3\xi_1 + 112\beta\xi_2 - 49\xi_3 = 0$$

becomes

$$27\xi_1 + 28\xi_2 + 49\xi_3 = 0.$$

The coordinates of the point i, therefore, are 27, 28, 49.

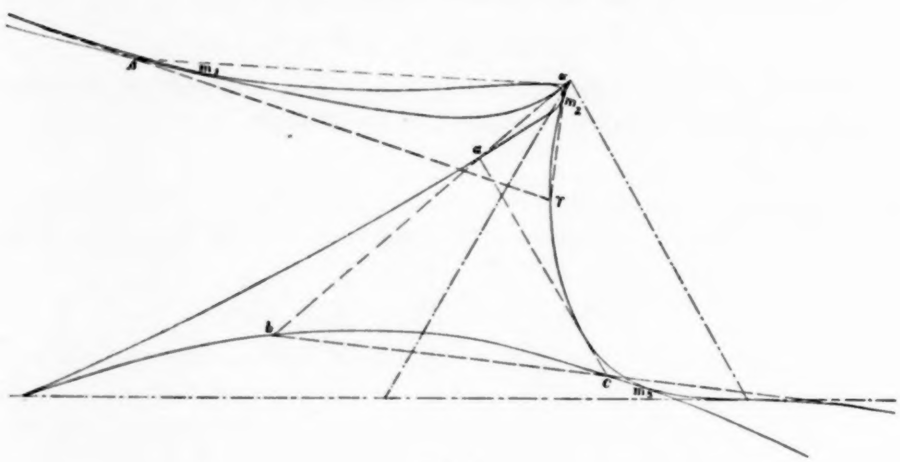


FIG. 4.

The contacts m_1, m_2, m_3 are given by the roots of the equation

$$14\lambda^3 + 30\lambda^2 - 12\lambda - 3 = 0,$$

and are approximately -2.46, $.51$, $-.18$. The intersection is given by the real root of

$$7\lambda^3 - 30\lambda^2 + 39\lambda - 24 = 0,$$

and lies between 2 and 3.

In figure 4 cubic f is drawn for $\beta = \frac{1}{2}$. Its cusp tangent is

$$x_1 - 2x_2 - 4x_3 = 0.$$

The cusp, given by $x_2 + 2x_3 = 0$, lies to the left of $0, 1, 0$, and at a distance equal to the base of the reference triangle from it.

The flex tangent

$$7x_2 + 8x_3 = 0$$

also cuts x_1 to the left of $0, 1, 0$ in the ratio $7 : -8$. The flex point is

$$216\xi_1 + 56\xi_2 + 49\xi_3 = 0.$$

The contacts m_1, m_2, m_3 are given by the equation

$$7\lambda^3 - 30\lambda^2 - 24\lambda + 12 = 0,$$

the roots of which corresponding to these points, respectively, are approximately $-.99$, 4.9, $.35$.

The equation for the intersections for this case is

$$7\lambda^3 + 60\lambda^2 + 156\lambda + 192 = 0,$$

the real root of which is approximately -5. The intersection, therefore, as we have already noticed, is outside of the triangle.

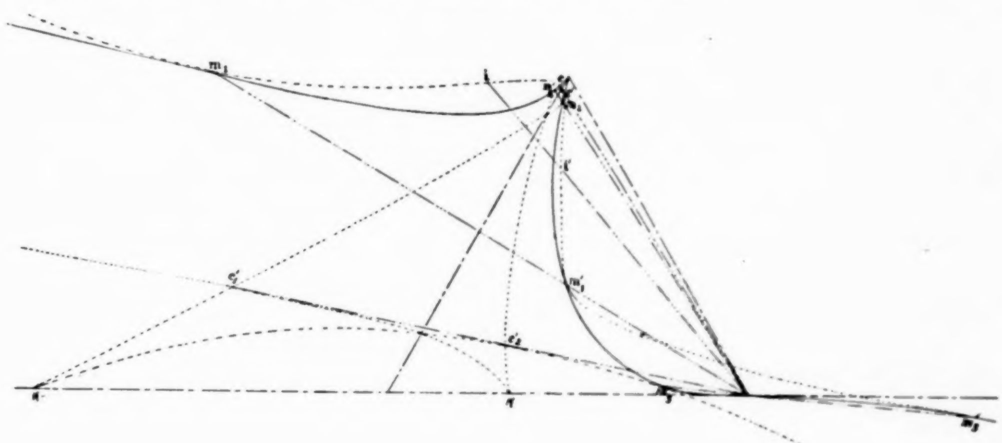


Fig. 5.

In each of the figures 3 and 4 $\alpha\beta\gamma$ is a triangle with vertices on ϕ and sides on f; and abc is one with vertices on f and sides on ϕ.

In figure 5 are shown the two positions of f for $\beta = \frac{1}{2}$ and $\beta = -\frac{1}{2}$. It shows the relation of the contacts of these with ϕ. It will be noticed that the contacts for $\beta = \frac{1}{2}$, say m_1, m_2, m_3, lie on lines with m_1', m_2', m_3',

respectively, through 0, 0, 1, m_1', m_2', m_3' being the contacts for $\beta = -\frac{1}{2}$. We have already noticed that these pairs are in an harmonic ratio with the point 0, 0, 1 and the point where x_3 cuts their join. This holds for the cusps κ and κ', the flexes i and i', the points of intersection c and c', and also the two points of intersection of the cubics themselves c_1' and c_2'. The other three points of intersection lie on x_3. In fact the curves are harmonic with 0, 0, 1 and x_3.

There is a definite relation between the cusp tangent and flex tangent for any position of f, or what is the same thing, the cusp and the point where the flex tangent cuts x_1. From the cusp tangent

$$x_1 - 8\beta^2 x_2 - 32\beta^3 x_3 = 0$$

we see that the cusp is given by

$$x_2 + 4\beta x_3 = 0.$$

Where the flex line cuts is given by

$$x_2' + \frac{16}{7}\beta x_3' = 0,$$

showing the one to one correspondence between the two. From these two equations we get the invariant relation or ratio

$$\frac{x_2}{x_3} = \frac{7}{4}\frac{x_2'}{x_3'} \qquad \text{or} \qquad \frac{x_2}{x_3} : \frac{x_2'}{x_3'} = \frac{7}{4}.$$

Starting, say, at the point 0, 0, 1 when $\beta = 0$, for which value the cusp lies at this point, and the flex tangent which is x_2 cuts x_1 at this point also, the cusp and the point of intersection of the flex tangent with x_1 moves to the left for negative values of β, the cusp at first having the greater velocity, let us say. The relative velocities change, and the cusp is overtaken and passed by the other point at 0, 1, 0 for $\beta = \infty$. They then traverse the external segment of the base for positive values of β, the relative velocities being interchanged with the above, the cusp overtaking the other point when 0, 0, 1 is reached, the ratio always being

$$\frac{x_2}{x_3} : \frac{x_2'}{x_3'} = \frac{7}{4}.$$

Johns Hopkins University,
Baltimore, Md.
April, 1906.

THE GROUPS GENERATED BY TWO OPERATORS SUCH THAT EACH IS TRANSFORMED INTO ITS INVERSE BY THE SQUARE OF THE OTHER

By G. A. Miller

Let s_1, s_2 be any two operators which satisfy the following conditions:

$$s_1^{-2}s_2s_1^2 = s_2^{-1}, \qquad s_2^{-2}s_1s_2^2 = s_1^{-1}.$$

Since s_1^2, s_2^2 transform each other into their inverses, they are either commutative or they generate the quaternion group;* in the former case either both of them are of order two and they generate the four group, or both of them are the identity. The hypothesis that one is of order 2 while the other is the identity leads to an absurdity in view of the given equations. Hence the orders of s_1, s_2 must have one of the following pairs of values: 1, 1; 1, 2; 2, 2; 4, 4; 8, 8. In the first two cases s_1, s_2 would generate the identity and the group of order two respectively. In the third cases they would generate the dihedral group. As these cases are practically trivial, we shall confine our attention in what follows to the groups generated by s_1, s_2 when their common order is either 4 or 8. The main results of this paragraph may be expressed as follows: *If two operators, neither of which is the identity, are such that each is transformed into its inverse by the square of the other, they must have the same order, and this common order is* 2, 4, *or* 8.

We begin with the case when s_1 is of order 8. The quaternion group $\{s_1^2, s_2^2\}$ generated by s_1^2, s_2^2 is invariant under the the group (G) generated by s_1, s_2. This fact results from the following equations:

$$s_1^{-1}s_2^2s_1 = s_1^{-1}s_2^2s_1s_2^{-2}s_2^2 = s_1^{-2}s_2^2, \qquad s_2^{-1}s_1^2s_2 = s_2^{-1}s_1^2s_2s_1^{-2}s_1^2 = s_2^{-2}s_1^2.$$

We shall now prove that the order of s_1s_2 is divisible by 3. This may be

* *Quarterly Journal of Mathematics*, vol. 37 (1906), p. 286.

done by means of the following equations:

$$s_1^{-1}s_2^{-1}s_2^2s_2s_1 = s_1^{-2}s_2^2,$$

$$s_1^{-1}s_2^{-1}s_1^{-2}s_2^2s_2s_1 = s_1^{-3}s_1^2s_2^{-1}s_1^{-2}s_2^3s_1 = s_1^{-3}s_2^4s_1 = s_1^{-3}s_1^4s_1 = s_1^2,$$

$$s_1^{-1}s_2^{-1}s_1^2s_2s_1 = s_1^{-1}s_2^{-2}s_1^3 = s_2^{-2}s_2^2s_1^{-1}s_2^{-2}s_1^3 = s_2^{-2}s_1^4 = s_2^2.$$

Since s_2s_1 transforms s_2^2, $s_1^{-2}s_2^2$, s_1^2 cyclically, its order is divisible by three. This proof is equally valid when the order of s_1 is 4, as we made no use of the fact that the order of s_1 was supposed to be 8 except when we stated that $\{s_1^2, s_2^2\}$ was the quaternion group. When s_1 is of order 4 this is the four-group. Hence we have the theorem: *If two operators whose orders exceed 2 are such that each is transformed into its inverse by the square of the other, the group generated by them involves operators of order* 3.

The group $\{s_1^2, s_2^2, s_2s_1\}$ is invariant under G, since $\{s_1^2, s_2^2\}$ is invariant under G and the following equations show that s_2s_1 is transformed into an operator of $\{s_1^2, s_2^2, s_2s_1\}$ by s_1 and s_2:

$$s_1^{-1}s_2s_1^2 = s_1^{-1}s_2^{-1} \cdot s_2^2s_1^2, \qquad s_2^{-1}s_2s_1s_2 = s_1s_2 = s_1^2s_1^{-1}s_2^{-1}s_2^2.$$

Hence G contains $\{s_1^2, s_2^2, s_2s_1\}$ as a sub-group of half its order. Since $(s_2s_1)^3$ is invariant under $\{s_1^2, s_2^2, s_2s_1\}$ and s_2s_1 transforms the operators of $\{s_1^2, s_2^2\}$ according to an operator of order 3, it follows that when s_2s_1 is of order 3, that is, when it has its least possible value, $\{s_1^2, s_2^2, s_2s_1\}$ is either the tetrahedral group or the group of order 24 which does not contain a sub-group of order 12, as the order of s_1 is 4 or 8. Hence $\{s_1^2, s_2^2, s_2s_1\}$ has an $(a, 1)$ isomorphism with the tetrahedral group whenever s_2s_1 is of order 3. Since the invariant operators of $\{s_1^2, s_2^2, s_2s_1\}$ are generated by s_1^4 and $(s_2s_1)^3$, it follows that $\{s_1^2, s_2^2, s_2s_1\}$ has an $(a, 1)$ isomorphism with the tetrahedral group regardless of the order of s_2s_1. The quotient group of G with respect to the sub-group generated by s_1^4, $(s_2s_1)^3$ is therefore one of the two groups of order 24 which contain the tetrahedral group. As this quotient group involves operators of order 4, it is the symmetric group of order 24. This proves the fundamental theorem: *If two operators whose orders exceed 2 are such that each is transformed into its inverse by the square of the other, they generate a group which has an* $(a, 1)$ *isomorphism with the symmetric group of order* 24. This evidently includes the theorem stated at the end of the preceding paragraph.

From the preceding paragraph it follows that the symmetric group of order 24 is the smallest possible group which may be generated by two operators whose orders exceed 2 and which are such that each is transformed into its inverse by the square of the other. This property may be regarded as a new definition of the symmetric group of order 24. Another definition which results from the preceding equations is this: Two operators of order 4 generate the symmetric group of order 24 provided their product is of order 3 and each of them is transformed into its inverse by the square of the other. Since the operators which correspond to the identity in the given $(a, 1)$ isomorphism between G and the symmetric group of order 24 generate an Abelian group it is clear that G contains exactly 4 Sylow subgroups of order 3^2 and that these are cyclic.

The operators of G which are invariant under $\}s_1^2, s_2^2, s_2s_1\{$ are generated by s_1^4 and $(s_1s_2)^3$, since they are generated by s_1^4 and $(s_2s_1)^3$. We proceed to prove that these operators are transformed into their inverses by the operators of G which are not in $\}s_1^2, s_2^2, s_2s_1\{$. This fact results from the following equations:

$$(s_1^2s_2)^{-1}s_1s_2s_1^{-2}s_1^2s_2 = s_2^{-1}s_1^{-1}s_2^2 = (s_1s_2)^{-1}s_2^2s_1s_2(s_1s_2)^{-1} = s_1^2(s_1s_2)^{-1},$$

$$(s_1s_2s_1^{-2})^3 = s_1s_2s_1^{-2}\cdot s_1s_2s_1^{-2}\cdot s_1s_2s_1^{-2} = (s_1s_2)^2s_2^2s_1s_2s_1^{-2} = (s_1s_2)^3.$$

Since $s_1^2s_2$ transforms $s_1s_2s_1^{-2}$ into its inverse and s_1^4 is invariant under G, the theorem under consideration is proved. Moreover, $(s_1^2s_2)^2 = s_1^2s_2s_1^2s_2 = s_1^4$. That is, G may be obtained by establishing a $(4, m)$ isomorphism between the symmetric group of order 24 and the dihedral group of order $6m$ whenever s_1 is of order 4.

It has been observed that s_1, s_2 may be so chosen that they generate the symmetric group of order 24 and that in all other cases the order of G is a multiple of 24. When s_1, s_2 are of order 4 they may be so selected that the order of G is any arbitrary multiple of 24. To do this we may replace the s_1, s_2 which generates the symmetric group of order 24 by s_1^1, s_2^1, where $s_1^1s_2^1$ are the products of s_1, s_2 respectively into two operators of order two which are independent of s_1, s_2 and whose product is the arbitrary multiple of 3. The same result is obtained by making the symmetric group of order 24 isomorphic with the dihedral group of order $6m$ in such a way that 4 operators of the former correspond to m operators of the latter, where m is arbitrary.

When s_1 is of order 8 the order of G is a multiple of 48, since the order of $\{s_1^2, s_2^2, s_1 s_2\}$ is a multiple of 24 in this case. It has been observed that the latter group is of order 24 and contains no sub-group of order 12 when $s_1 s_2$ is of order 3. When $s_1 s_2$ satisfies this condition G is one of the four groups of order 48 which contain this group of order 24*. As it has a (2, 1) isomorphism with the symmetric group of order 24 and as $s_1^2 s_2$ is of order 4 while s_1 is of order 8, it contains 12 operators of each of the orders 4 and 8 besides those of $\{s_1^2, s_2^2, s_1 s_2\}$. These conditions determine this group of order 48 completely and prove that it is the one known as G_{52}.* *The smallest possible group which is generated by two operators of order 8 such that each is transformed into its inverse by the square of the other is the group of order 48 known as G_{52}*, and every other group which can be generated by two such operators has an $(a, 1)$ isomorphism with this group. The method used in the preceding paragraph may be employed to show that s_1, s_2 can be so chosen that a has any arbitrary value.

The group of order 48 whose properties were considered in the preceding paragraph may be represented as a substitution group of degree 16 in the following manner:

$$s_1 = ac'eg'bd'fh' \cdot a'de'hb'cf'g, \qquad s_2 = ah'cf'bg'de' \cdot a'gc'eb'hd'f,$$

$$s_1^2 = aebf \cdot cdgh \cdot a'e'b'f' \cdot c'g'd'h', \qquad s_2^2 = acbd \cdot ehfg \cdot a'c'b'd' \cdot e'h'f'g',$$

$$s_2^{-2} s_1 s_2^2 = ah'fd'bg'ec' \cdot a'gf'cb'he'd = s_1^{-1}, \qquad s_1^{-2} s_2 s_1^2 = ae'dg'bf'ch' \cdot a'fd'hb'ec'g = s_2^{-1},$$

$$s_1 s_2 = aed \cdot bfc \cdot a'e'd' \cdot b'f'c'.$$

Since s_1^2, s_2^2 transform each other into their inverses they generate the quaternion group which is transformed into itself by $s_1 s_2$. Hence the remaining substitution of this group of order 48 could easily be written out. When s_1 is of order 4, G is completely determined by the order of $s_1 s_2$ and the fact that each of the operators s_1, s_2 is transformed into its inverse by the square of the other. This is, however, not the case when s_1 is of order 8. In this case it is necessary to specify whether the operator of order 2 generated by $s_1 s_2$ is in $\{s_1^2, s_2^2\}$ or not. For instance, the two substitutions

$$s_1 = ac'eg'bd'fh' \cdot a'de'hb'cf'g, \qquad s_2 = ag'ce'bh'df' \cdot a'hc'fb'gde'$$

are such that each is transformed into its inverse by the square of the other

---

* *Quarterly Journal of Mathematics*, vol. 30 (1899), p. 258.

but their product is of order 6. Since the cube of this product is in $\{s_1^2, s_2^2\}$, they generate the same group of order 48 as the two substitutions given above.

As a rule the most useful groups are those which admit the simplest definitions by means of the laws of combination of their symbols. The groups of genus zero are the most remarkable from this standpoint. The systems whose fundamental properties were considered above appear also to be remarkable from the standpoint of simplicity of definition, and their close contact with the symmetric group of order 24 adds interest. The new definition of the latter group which flows from a study of these systems seems worthy of notice,— not so much on account of its newness as on account of its simplicity. The category of groups considered above suggests the more general question in reference to the groups which are generated by n operators ($n > 2$) which are such that each is transformed into its inverse by the squares of all the others. While it is known that the squares of these operators would generate either a Hamiltonian group or an Abelian group of order 2^a and of type $(1, 1, 1, \cdots)$* which would be an invariant sub-group and that the order of the group which they generate is divisible by 3 whenever the order of at least one of the generating operators exceeds 2, yet the category of these groups includes so many different types as to make it improbable that much progress can be made along this line. Since every possible symmetric group can be generated by three operators of order 2,† it follows that when $n > 2$ every possible symmetric group could be a quotient group of some groups whose generators would satisfy the given conditions.

University of Illinois,
December, 1906.

* *Quarterly Journal of Mathematics*, vol. 37 (1906), p. 286.
† Cf. *Bulletin of the American Mathematical Society*, vol. 7 (1901), p. 426.

ON THE CLASSIFICATION OF PLANE ALGEBRAIC CURVES POSSESSING FOURFOLD SYMMETRY WITH RESPECT TO A POINT*

By R. D. Carmichael

A plane figure is said to be symmetrical with respect to a point, called the center of symmetry, if it coincides with its original position after being turned through 180° in its plane about that center. This we may call twofold symmetry.

Generally, we shall say that a plane figure has r-fold symmetry with respect to a point, called the center of symmetry, if it coincides with its original position after being turned through $\frac{360^\circ}{r}$ in its plane about that center.

The object of this paper is the classification of irreducible plane algebraic curves possessing fourfold symmetry with respect to a point. The work will be carried out by showing that (excluding the case of curves consisting solely of isolated points) there are no such curves of odd degree, and that there are two classes of such curves of even degree.

If the origin of rectangular Cartesian coordinates is taken at the center of symmetry, and if (α, β) is a point on the curve, then also is $(-\alpha, -\beta)$, as may be readily seen from the above definition of fourfold symmetry. Then if these two sets of values are put for x, y in the equation of the curve there will result two equations which differ only in respect to the signs of the terms of odd degree in α, β, all such signs being different in the two cases. Take the sum of these two equations if their degree is odd: and their difference if their degree is even: in each case the resulting equation will be one of degree $\leqq n-1$, n being the degree of the original equation. This is inconsistent with the assumption of a locus of the nth order unless the degree of this last equation is zero; that is, unless the equation vanishes. For, suppose the contrary; we should then have a locus of order less than n containing all the points (α, β) of an nth degree locus; and this is impossible since the curve is supposed irreducible. This leads readily to the following preliminary theorem:

* Read before the American Mathematical Society, September 5, 1907.

The equation in rectangular Cartesian coordinates of a plane algebraic curve of order n, possessing fourfold symmetry with respect to a point taken as the origin, has only terms of odd degree in x, y when n is odd, and only terms of even degree in x, y when n is even.

This equation may be written

$$(1) \qquad \sum a_{ts} x^t y^s = 0,$$

where t and s each ranges from 0 to n subject to the condition

$$t + s = 1, 3, 5, \cdots, n \quad \text{when } n \text{ is odd},$$

$$t + s = 0, 2, 4, \cdots, n \quad \text{when } n \text{ is even},$$

and where a_{ts} are real constants so far taken independent of each other. We proceed now to find certain relations which exist among them.

For this purpose it will be found convenient to transform the equation to polar coordinates, the origin remaining the same and the angle θ being measured from the x-axis. The equation then is

$$(2) \qquad \sum a_{ts} \cos^t\theta \sin^s\theta \, \rho^{t+s} = 0,$$

where t and s are subject to the same conditions as before. Now, if $90° + \theta$ is put for θ, equation (2) becomes

$$(3) \qquad \sum (-1)^t a_{ts} \sin^t\theta \cos^s\theta \, \rho^{t+s} = 0.$$

In order that the loci should possess the defined fourfold symmetry with respect to the origin it is both necessary and sufficient that the same values of ρ satisfy both equations (2) and (3) when θ has any fixed value. Therefore the coefficients of like powers of ρ can differ only by some constant factor m. Therefore we must have

$$(4) \qquad \sum a_{ts} \cos^t\theta \sin^s\theta = m \sum (-1)^t a_{ts} \sin^t\theta \cos^s\theta,$$

where the summation is extended to all permissible values of t and s such that $t + s$ has any fixed value.

We assume that the locus does not consist entirely of isolated points.

Then there are an infinite number of values of θ for which $\cos\theta$ is not zero in every case for which $n > 1$; and evidently when $n = 1$ the fourfold symmetry does not exist. Then transposing in (4) and dividing by $\cos^{t+s}\theta$, we have

$$(5) \qquad \sum a_{ts} \tan^s\theta - m \sum (-1)^t a_{ts} \tan^t\theta = 0,$$

$t + s$ being fixed in value. Now, (5) is an algebraic equation in $\tan\theta$ of degree not exceeding n; but it must be satisfied by an infinite number of values of θ. Therefore its first member vanishes identically; and hence the coefficient of any power of $\tan\theta$ must vanish. That is, we must have

$$(6) \qquad a_{st} - m(-1)^t a_{ts} = 0.$$

Interchanging t and s we arrive at the equation

$$a_{ts} - m(-1)^s a_{st} = 0;$$

whence

$$m(-1)^t a_{ts} - m^2(-1)^{t+s} a_{st} = 0.$$

Combining with (6), we have

$$a_{st} - m^2(-1)^{t+s} a_{st} = 0;$$

or

$$1 - m^2(-1)^{t+s} = 0, \qquad \text{or} \qquad a_{st} = a_{ts} = 0.$$

In the latter case, $a_{st} = a_{ts}$, equation (1) vanishes. In the former case, if n is odd so is $t + s$; and therefore we have

$$m = \pm\sqrt{-1};$$

that is, two real coefficients differ by an imaginary factor, which is impossible. Therefore we have the following theorem:

Excluding curves consisting of isolated points only, there exist no loci of odd order possessing fourfold symmetry with respect to a point.

If n is even so is $t + s$; then we have

$$m = +1, \quad -1.$$

Hence from (6) it follows readily that there are to be considered the two cases

$$(7) \qquad a_{st} = (-1)^t a_{ts}; \qquad \text{and} \qquad a_{st} = -(-1)^t a_{ts}.$$

That is, these are conditions both necessary and sufficient for the existence of the defined symmetry.

By the aid respectively of the first and the second conditions in equations (7) we may transform equation (1) into the following:

$$\sum [a_{ts}(x^t y^s + (-1)^t x^s y^t)] = 0, \qquad t \geqq s;$$

$$\sum [a_{ts}(x^t y^s - (-1)^t x^s y^t)] = 0, \qquad t \geqq s.$$

The condition $t \geqq s$ may evidently be replaced by $t = s, s + 1, \cdots, n$. Putting these results with the preceding ones we have the theorem:

Excluding curves consisting of isolated points only, the only loci of the nth order possessing fourfold symmetry with respect to a point taken as the origin are those whose rectangular Cartesian equations are of one of the forms

$$\sum [a^{ts}(x^t y^s + (-1)^t x^s y^t)] = 0,$$

$$\sum [a_{ts}(x^t y^s - (-1)^t x^s y^t)] = 0,$$

where n must be even and

$$s = 0, 1, 2, \cdots, n,$$
$$t = s, s + 1, \cdots, n;$$

subject to the condition that

$$t + s = 0, 2, 4, \cdots, n.$$

Alabama Presbyterian College,
Anniston, Ala.,
June, 1907.

A SECOND INVERSE PROBLEM OF THE CALCULUS OF VARIATIONS*

By Carl Eben Stromquist

The simplest problem in the calculus of variations is the following: Given an integral

$$(1)\qquad I=\int_{x_1}^{x_2} g(x, y, p)\,dx, \qquad \text{where } p=\frac{dy}{dx},$$

to find a curve $y = y(x)$ which joins two given fixed points and renders this integral a minimum. A first necessary condition for a minimum is that $y = y(x)$ must satisfy the Euler differential equation

$$(2)\qquad y''g_{pp} + pg_{py} + g_{px} - g_y = 0,$$

where $y'' = d^2y/dx^2$ and the subscripts denote partial differentiation. The two-parameter family of curves

$$(3)\qquad y = y(x, \alpha, \beta)$$

which the differential equation (2) yields is called the family of extremals of the integral I.

The so-called inverse problem of the calculus of variations arises when a two-parameter family of curves

$$y = y(x, \alpha, \beta),$$

and hence their differential equation

$$(4)\qquad y'' = \phi(x, y, p),$$

* Read before the American Mathematical Society, April 27, 1907.

are given and one seeks to determine the integrand $g(x, y, p)$ so that the given family of curves (3) will be extremals for the integral

$$I = \int_{x_1}^{x_2} g(x, y, p)\,dx.$$

The inverse problem in this general form has been discussed by Darboux.* If the extremals are to be the solutions of equation (4), the value of y'' substituted in (2) must reduce this equation to an identity in x, y, p and hence the equation

$$\phi(x, y, p) g_{pp} + p g_{py} + g_{px} - g_y = 0$$

will be a differential equation for g. As there are an infinite number of values of g which will satisfy this differential equation, the inverse problem has in general an infinite number of solutions. Completely to determine g some further conditions must therefore be imposed. It has been shown† that in the cases where straight lines or circles with centers on the x-axis are extremals, if the further condition be imposed that extremals are perpendicular to their transversals, then the function g is completely determined aside from a constant factor. The condition of transversality is a relation between the variables x, y, p, P, where p is the slope of the extremal at the point (x, y), and P the slope of the curve cutting the extremal transversally at that point. The equation expressing this relation is in fact ‡

$$g(x, y, p) + (P - p) g_p(x, y, p) = 0, \tag{5}$$

which when solved for P takes the form§

$$P = \theta(x, y, p). \tag{6}$$

We have seen above that when the integral (1) is given the extremals (3) can always be determined, or inversely when the family of curves (3) is given then an integral with these as extremals can always be found.

---

* Darboux, *Leçons*, vol. 3, p. 53. See also Bolza, *Lectures on the Calculus of Variations*, p. 30.

† Stromquist, *Transactions Amer. Math. Society*, vol. 7 (1906), p. 181.

‡ Bolza, loc. cit., pp. 36, 106, 172. Geometrically stated, two transversals to the same set of extremals are two curves which intercept on the extremals arcs along which the integral under consideration has a constant value.

§ It is supposed in what follows that the function θ and its first partial derivatives exist and are continuous in x, y, p and that $p - \theta \neq 0$.

A second inverse problem in connection with the integral (1) is the following: If a relation of the form (6) is given, i. e., the relation between the direction whose slope is p and the direction transverse to it at each point of the plane, then will it always be possible to determine an integral (1) whose extremals and transversals satisfy the given relation? It will be shown in §1 that a function g giving rise to the relation (6) can always be determined, and the further conditions which insure that it is uniquely determined will be derived. In §2 the results of §1 will be applied to some examples.

1. Determination of the integral corresponding to a given relation of transversality. If the integral (1) is to have the equation (6) as its relation of transversality then it is evident from (5) that the integrand g must be a solution of the equation

$$g(x, y, p) + \{\theta(x, y, p) - p\} g_p(x, y, p) = 0$$

for all values of x, y, p. This is a linear differential equation for g which is easily integrated. The solution has the form

$$g = \psi(x, y)e^A, \tag{7}$$

where

$$A = \int_0^p \frac{dp}{p - \theta(x, y, p)},$$

and $\psi(x, y)$ is an arbitrary function of x and y. We thus have as a first result: *If an integral*

$$I = \int_{x_1}^{x_2} g(x, y, p)dx$$

has the equation

$$P = \theta(x, y, p)$$

as its relation of transversality, then the integrand g must be of the form

$$g = \psi(x, y)e^A,$$

where

$$A = \int_0^p \frac{dp}{p - \theta(x, y, p)},$$

and $\psi(x, y)$ is an arbitrary function of x and y.

In order to determine the arbitrary function $\psi(x, y)$ which enters into the expression for g, suppose that it be further required that a one-parameter family of curves

$$y = y(x, \gamma) \tag{8}$$

be given which shall be extremals for the integral I. It is well known that if along an arc of one of these extremals the derivative $\partial y/\partial\gamma$ is different from zero then a neighborhood of that arc can be found which is simply covered by the extremals (8) and which constitutes a so-called field.* Let us suppose that such a field has been found. Then at each point (x, y) of the field the equation (8) has a solution

$$\gamma = \gamma(x, y),$$

which gives the value of γ for the extremal of the family which passes through that point. The slope of the extremal at the point (x, y) has the value

$$p = \frac{dy(x, \gamma)}{dx} = \frac{dy(x, \gamma(x, y))}{dx} = p(x, y), \tag{9}$$

and the second derivative of y is

$$y'' = p_x + pp_y. \tag{10}$$

Hence, if the given relation (6) is to exist between the directions of the extremals and transversals and if further $y = y(x, \gamma)$ is to be a one-parameter family of extremals for the integral (1), then the Euler differential equation (2) must be satisfied identically when the value of g given by (7) and the value of y'' given by (10) are substituted. The Euler equation then becomes

$$\frac{(p_x + pp_y)\theta_p + \theta_x + p\theta_y}{p - \theta} + \int_0^p \frac{\theta_x dp}{(p-\theta)^2} + \theta\int_0^p \frac{\theta_y dp}{(p-\theta)^2} + \frac{1}{\psi}\left(\psi_x + \theta\psi_y\right) = 0, \tag{11}$$

where the value $p(x, y)$ from equation (9) is to be substituted for p. Equation (11) is a partial differential equation of the first order for the function ψ. It can be written in the form

$$\psi_x + \theta\psi_y + K(x, y)\psi = 0, \tag{12}$$

* Bolza, loc. cit., §19; Osgood, ANNALS OF MATHEMATICS, ser. 2, vol. 2, p. 112; Goursat, *Cours d'Analyse*, vol. 2, §451. See also these references for conditions of continuity etc. that must be satisfied by $y(x, \gamma)$.

where

$$K(x, y) = \frac{(p_x + pp_y)\theta_p + \theta_x + p\theta_y}{p - \theta} + \int_0^p \frac{\theta_x dp}{(p-\theta)^2} + \theta \int_0^p \frac{\theta_y dp}{(p-\theta)^2}.$$

It is evident that this equation still does not uniquely determine the function ψ.

In order to integrate equation (12) we have the auxiliary set of equations*

$$\frac{dx}{1} = \frac{dy}{\theta} = \frac{d\psi}{-K(x, y)\psi}. \tag{13}$$

Let the solution of

$$dx = \frac{dy}{\theta}$$

be taken in the form

$$u(x, y) = c. \tag{14}$$

If this is solved for y in terms of x and c giving $y = \eta(x, c)$ then the solution of

$$\frac{d\psi}{-K(x, \eta)\psi} = \frac{dx}{1}$$

will be of the form

$$\psi \cdot v(x, y) = c_1, \tag{14'}$$

where v is a particular solution of the equation

$$\frac{dv}{dx} - K(x, \eta)v = 0,$$

and is found by solving this equation regarding c as a constant and afterwards replacing c by the function $u(x, y)$. Then the most general solution of equation (12) is

$$\psi \cdot v(x, y) = W(u), \tag{15}$$

* Darboux, loc. cit., vol. 3, p. 55; Jordan, *Cours d'Analyse*, vol. 3, §242.

where W is an arbitrary function of u. Equation (7) now reduces to

$$g(x, y, p) = \frac{W(u)}{v(x, y)} \cdot e^{A}, \tag{16}$$

where
$$A = \int_0^p \frac{dp}{p - \theta(x, y, p)}.$$

Here u and v are known functions of x and y but W is an arbitrary function of u.

It is to be observed that any curve $u = c$ is a transversal in the field of extremals (8) since it is the solution of $dy/dx = \theta$. If the value of I is given along any curve of the field which is not one of these transversals or tangent to one of them, then the function $W(u)$, and consequently $g(x, y, p)$, is uniquely determined. For along such a curve, $y = y(x)$, the derivative

$$\frac{du}{dx} = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y}\frac{dy}{dx}$$

is different from zero so that u is an always increasing (or always decreasing) function of x along the curve, and conversely x is an always increasing (or always decreasing) function $x = x(u)$ of u. Since now I is an assigned function of x along the curve $y = y(x)$ it follows that $dI/dx = g$ in equation (16) is uniquely determined. Hence if we substitute in the right hand member of equation (16) $y = y(x)$ and $p = y'(x)$, and then replace x by the function $x = x(u)$, the result will be an identity in u which uniquely determines the function W. We have then attained the following result:

If besides the relation of transversality (6) it is also prescribed that the integral I shall have a one-parameter family of curves $y = y(x, \gamma)$ (8) arbitrarily chosen, as extremals, then the function $\psi(x, y)$ in equation (7) must satisfy a partial differential equation (12). The value of $\psi(x, y)$ is completely determined if in a field composed of extremals (8) the value of the integral I as a function of x is given along any curve which is not tangent at any point to a transversal of the field.

2. Examples. In this section the above results will be applied to some examples.

Example 1. Consider the case where the extremals and transversals are perpendicular, i. e., where

$$P = \theta = -\frac{1}{p}. \tag{17}$$

It follows readily from equation (7) that then g must have the form

$$g = \psi(x, y)\sqrt{1 + p^2}, \tag{18}$$

and equation (11), from which ψ is to be found, becomes

$$(p_x + pp_y)\psi + (1 + p^2)(p\psi_x - \psi_y) = 0. \tag{19}$$

Example 1*a*. Suppose it be further required that the one-parameter family of straight lines through the fixed point $P_0(x_0, y_0)$, viz.

$$y - y_0 = m(x - x_0),$$

be extremals. Since then

$$p = \frac{dy}{dx} = m = \frac{y - y_0}{x - x_0},$$

it follows that

$$p_x + pp_y = 0,$$

and hence equation (19) becomes

$$(y - y_0)\psi_x - (x - x_0)\psi_y = 0. \tag{20}$$

The auxiliary set of equations to be integrated in this case is

$$\frac{dx}{1} = \frac{dy}{-\dfrac{x - x_0}{y - y_0}} = \frac{d\psi}{0},$$

from which are obtained the following results—corresponding to equations (14) and (14′):—

$$u = (x - x_0)^2 + (y - y_0)^2 = c^2, \tag{21}$$

$$\psi = c_1;$$

hence the general solution of equation (19)—corresponding to equation (15)—is

$$\psi = W\{(x - x_0)^2 + (y - y_0)^2\}.$$

In this case, therefore, g must have the form

$$g(x, y, p) = W(u)\sqrt{1 + p^2}.$$

The transversals, as is seen from equation (21), are a family of circles with centres at the given point (x_0, y_0). Suppose that along some curve L which passes through (x_0, y_0) and is not tangent at any point to one of these transversals the value of the integral

$$\int_{x_1}^{x} g\,dx = \int_{x_1}^{x} W(u)\sqrt{1 + p^2}\,dx$$

be equal to the integral $\int_{x_1}^{x} \sqrt{1 + p^2}\,dx$. It then follows by differentiating for x that for all values of u along the given curve, $W(u) = 1$. Hence g is completely determined and must be of the form $g = \sqrt{1 + p^2}$ at least for all points in a circle about (x_0, y_0) as a centre and with the largest value of u on L as a radius.

These results have an interesting application in the theory of plane geometry. We may assume that the elements of our geometry are the number pairs (x, y) and may then define length * along any curve to be given by an integral of the form

$$I = \int g(x, y, p)\,dx. \tag{22}$$

The nature of the geometry is then determined by the character of the function g and in the special case when g has the form $\sqrt{1 + p^2}$ we get the ordinary theory of Euclidean geometry. The extremals for the integral (22) play the role of the straight lines in the ordinary geometry and the circles about a point P_0 correspond to the transversals to the one-parameter family of extremals through P_0. The condition of transversality (6) is then a relation between

* This is practically the standpoint of Hamel in his dissertation, Göttingen, 1901. See also *Transactions Amer. Math. Society*, vol. 7 (1906), p. 175, third foot-note.

the directions of a shortest line and a circle. With these geometrical notions, the above results can be stated as follows:

If the length of an arc of a curve from (x_1, y_1) to (x, y) be defined as the value of the integral $I = \int_{x_1}^{x} g(x, y, p)\, dx$, and if circles are required to be perpendicular to their radii (i. e., transversals perpendicular to extremals), then g must be of the form $g = \psi(x, y)\sqrt{1 + p^2}$. If it be further true that there is a single point P_0 such that the one-parameter family of shortest lines through P_0 are all straight lines and if the integral I takes the same values as the Euclidean length integral along a single curve through the point P_0 and not tangent to any of the circles about P_0, then it follows that I must be everywhere equal to $\int_{x_1}^{x} \sqrt{1 + p^2}\, dx$.

Example 1*b*. If it be required that the one-parameter family of circles with centres on the x-axis and passing through the point (x_0, y_0), viz.

$$x^2 - x_0^2 + y^2 - y_0^2 - 2a(x - x_0) = 0, \tag{23}$$

be extremals, then equation (19) reduces to

$$\{(x - x_0)^2 - (y^2 - y_0^2)\}\psi_x + 2y(x - x_0)\psi_y + 2(x - x_0)\psi = 0. \tag{24}$$

The equations corresponding to equations (14) and (14′) are in this case

$$u = \frac{(x - x_0)^2 + y^2 + y_0^2}{y} = c,$$

$$y \cdot \psi = c_1.$$

Hence the most general solution of equation (24) is

$$\psi = \frac{1}{y} W\left\{\frac{(x - x_0)^2 + y^2 + y_0^2}{y}\right\},$$

and therefore g must have the form

$$g = \frac{1}{y} W(u)\sqrt{1 + p^2}.$$

In a manner analogous to that of the preceding example it can now be shown that, if we further require that along a single curve through the point (x_0, y_0) and not tangent to a transversal the value of

$$\int_{x_1}^{x} g\,dx = \int_{x_1}^{x} \frac{1}{y} W(u)\sqrt{1+p^2}\,dx$$

be equal to

$$\int_{x_1}^{x} \frac{1}{y}\sqrt{1+p^2}\,dx,$$

then g is completely determined and the integral I must have the form

$$\int_{x_1}^{x} g\,dx = \int_{x_1}^{x} \frac{1}{y}\sqrt{1+p^2}\,dx.$$

This is the form of the length integral obtained by a conformal transformation of the pseudo-sphere upon the plane.* We have therefore obtained the following result:

If it be required that circles are perpendicular to their radii (i. e., transversals perpendicular to extremals) and also that a one-parameter family of extremals through P_0 be the circles through this point and with centres on the x-axis, then it follows that the integral I is uniquely determined and has the form

$$I = \int_{x_1}^{x} \frac{1}{y}\sqrt{1+p^2}\,dx,$$

i. e., has the form of the length integral obtained by a conformal transformation of the pseudo-sphere upon the plane, provided that the integral has this value along a single curve through P_0 which is not tangent to any of the circles (transversals) about P_0.

Example 2. The case where a linear fractional relation,

$$P = \theta(x, y, p) = \frac{ap+b}{cp+d}, \tag{25}$$

exists between P and p naturally suggests itself. Here a, b, c, d are functions of x and y. It follows from equation (5) that g in this case must be a

* See *Transactions*, loc. cit., p. 181.

solution of the linear differential equation

$$\frac{dg}{g} = \frac{(cp + d)dp}{cp^2 + (d - a)p - b}$$

for all values of x and y. This is readily integrated and the solution has the form

$$(26) \qquad g(x, y, p) = \psi(x, y)\sqrt{cp^2 + (d - a)p - b} \cdot e^{\frac{a+d}{2}H},$$

where

$$H = \frac{2}{\sqrt{-q}} \tan^{-1} \frac{2cp + (d - a)}{\sqrt{-q}}$$

if $cp^2 + (d - a)p - b$ is a definite quadratic form, or

$$H = \frac{1}{\sqrt{q}} \log \frac{2cp + d - a - \sqrt{q}}{2cp + d - a + \sqrt{q}}$$

if $cp^2 + (d - a)p - b$ is an indefinite quadratic form. Here $q = (d - a)^2 + 4bc$. If $cp^2 + (d - a)p - b$ is a semi-definite quadratic form, then

$$g = \psi(x, y)(2cp + d - a)e^{-\frac{d+a}{2cp+d-a}}.$$

Example 3. It is well known from the theory of surfaces* that if the line element of a surface be of the form

$$g\,dx = \sqrt{E + 2Fp + Gp^2}\,dx,$$

where x and y are the arguments of E, F, G in place of the ordinary u and v, then the condition that a family of curves

$$\phi_2(x, y) = \text{constant}$$

shall be the orthogonal trajectories of a given family of curves

$$\phi_1(x, y) = \text{constant}$$

is that

$$P = \frac{-Fp - E}{Gp + F},$$

* Bianchi, *Vorlesungen über Differentialgeometrie*, p. 66.

where

$$p = -\frac{\dfrac{\partial \phi_1}{\partial x}}{\dfrac{\partial \phi_1}{\partial y}} \qquad \text{and} \qquad P = -\frac{\dfrac{\partial \phi_2}{\partial x}}{\dfrac{\partial \phi_2}{\partial y}}.$$

From equation (26) of the preceding example it follows at once that *if it be required that the condition of orthogonality of two curves on a surface be*

$$P = \frac{-Fp - E}{Gp + F},$$

where E, F, G are functions of the parameters x and y of the surface, then the line element of the surface must have the form

$$g\,dx = \psi(x, y)\sqrt{E + 2Fp + Gp^2}\,dx.$$

Example 4. As a further special case of example 2, it follows readily that if extremals and transversals cut at a constant angle $\beta = \tan^{-1} k$, i. e., if

$$\theta = \frac{p - k}{kp + 1},$$

then the integrand g must have the form

$$g = \psi(x, y)e^{\frac{1}{k}\tan^{-1}p}\sqrt{1 + p^2}.$$

In this case if a one-parameter family of straight lines through the origin, $y = mx$, be extremals then it follows that the transversals will be the spirals

$$\sqrt{x^2 + y^2} \cdot e^{\frac{1}{k}\tan^{-1}\frac{y}{x}} = c \qquad \text{or} \qquad \rho = ce^{-\frac{\theta}{k}}.$$

If the circles $x^2 + y^2 - 2ax = 0$ are extremals the transversals are the circles

$$(x - c)^2 + (y - ck)^2 = c^2(1 + k^2).$$

If the circles $x^2 + y^2 = a^2$ are extremals the transversals are the spirals

$$\sqrt{x^2 + y^2}\, e^{-k\tan^{-1}\frac{y}{x}} = c \qquad \text{or} \qquad \rho = ce^{k\theta}.$$

PRINCETON, N. J.,
JUNE, 1907.

THE CONTINUOUS PLANE MOTION OF A LIQUID BOUNDED BY TWO RIGHT LINES

By Henry C. Wolff

The following paper gives a discussion of the general equation transforming an infinite strip, of definite width, conformally upon the entire plane: the boundaries of the strip become any two non-intersecting lines, each extending to infinity in one direction. This general function, of which the hyperbolic sine,* the well known function of Helmholtz,† and the function recently discussed by Harris,‡ are special cases, is expressed below by equation (10).

Let us consider a property of the special functions which led to the discovery of the general equation. From the equation

$$Z = \pi \sinh \frac{z}{2}$$

we have

$$\frac{dZ}{dz} = \zeta = \frac{\pi}{4}\left[e^{z/2} + e^{-z/2}\right].$$

Factoring the right-hand side as the difference of two perfect squares we obtain

$$\zeta = \frac{\pi}{4}\left[e^{z/4} - ie^{-z/4}\right]\left[e^{z/4} + ie^{-z/4}\right].$$

Let us write

$$\frac{dZ'}{dz} = \zeta' = \frac{\pi}{4}\left[e^{z/4} - ie^{-z/4}\right],$$

thus omitting the last factor of ζ.

Integrating we have

$$Z' = \pi\left[e^{z/4} + ie^{-z/4}\right]. \tag{1}$$

* A. R. Forsythe, *Theory of Functions of a Complex Variable*, Cambridge, 1893, p. 503.

† H. von Helmholtz, *Wissenschaftliche Abhandlungen*, vol. 1, p. 154.

‡ R. A. Harris, On Two-dimensional Fluid Motion through Spouts composed of Two Plane Walls, Annals of Mathematics, ser. 2, vol. 2, p. 73; 1902.

Separating the real and imaginary parts we have

$$X' = \pi\left[e^{x/4}\cos\frac{y}{4} + e^{-x/4}\sin\frac{y}{4}\right],$$

$$Y' = \pi\left[e^{x/4}\sin\frac{y}{4} + e^{-x/4}\cos\frac{y}{4}\right].$$

When $y = \pi$, let X and Y be represented by $X(\pi)$ and $Y(\pi)$ respectively. Then

$$X'(\pi) = \frac{\pi}{\sqrt{2}}\left[e^{x/4} + e^{-x/4}\right],$$

$$Y'(\pi) = \frac{\pi}{\sqrt{2}}\left[e^{x/4} + e^{-x/4}\right],$$

or

$$X'(\pi) = Y'(\pi),$$

the equation of a straight line in the Z'-plane. Here $X'(\pi)$ has its minimum value when $x = 0$.

When $y = -\pi$ we have

$$X'(-\pi) = \frac{\pi}{\sqrt{2}}\left[e^{x/4} - e^{-x/4}\right],$$

$$Y'(-\pi) = \frac{-\pi}{\sqrt{2}}\left[e^{x/4} - e^{-x/4}\right];$$

or

$$X'(-\pi) = -Y'(-\pi),$$

the equation of a straight line in the Z'-plane. Here $X'(-\pi)$ has neither a maximum nor a minimum value.

When $y = -3\pi$ we get

$$X'(-3\pi) = -\frac{\pi}{\sqrt{2}}\left[e^{x/4} + e^{-x/4}\right],$$

$$Y'(-3\pi) = -\frac{\pi}{\sqrt{2}}\left[e^{x/4} + e^{-x/4}\right];$$

or

$$X'(-3\pi) = Y'(-3\pi).$$

Here $X'(-3\pi)$ has its maximum value when $x = 0$.

Thus a strip of width 4π between the lines $y = \pi$ and $y = -3\pi$ on the z-plane is represented conformally upon the entire Z'-plane. The line $y = \pi$ is bent at the point $x = 0$, and the parts on either side are folded together form-

ing one half line in the Z'-plane. The line $y = -3\pi$ is bent at the point $x = 0$, and the two parts are folded together forming the second half of the same line in the Z'-plane. The line $y = -\pi$ remains unbroken. Thus we see that the transformations

$$Z = \pi \sinh \frac{z}{2} \quad \text{and} \quad Z' = \pi[e^{z/4} + ie^{-z/4}]$$

are similar, but the strip transformed by the first equation is one-half as wide as, and coincides with the upper half of, the strip transformed by the second equation. In fact, with this object in view, the latter equation could have been built up directly from the former.

Let us now, considering the second variable factor of ζ, write

$$\frac{dZ''}{dz} = \zeta'' = \frac{\pi}{4}\left[e^{z/4} + ie^{-z/4}\right].$$

Integrating, we have

$$Z'' = \pi[e^{z/4} - ie^{-z/4}]. \tag{2}$$

Discussing this in the same way that we discussed Z' we find that it represents the strip between the lines $y = -\pi$ and $y = +3\pi$ in the z-plane conformally upon the entire Z''-plane.

Now let us note how the functions derived from these three ζ's transform the lines $y = \pi$ and $y = -\pi$ on the z-plane. That derived from ζ' leaves the line $y = -\pi$ unbent, while it bends the line $y = \pi$ at the point $x = 0$. That derived from ζ'' does just the opposite; it leaves the line $y = \pi$ unbent and bends the line $y = -\pi$ at the point $x = 0$. The function corresponding to ζ bends both of these lines at the point $x = 0$. We may then look upon the functions, corresponding to the two factors of ζ, as each doing a half of the transformation performed by the function from which they were derived. It is apparent that if the z in ζ' were first changed to $z + a$ the transformation represented by equation (1) would bend the line $y = \pi$ at the point $x = -a$ instead of at the point $x = 0$; and if in equation (2) the z were replaced by $z - a$ the transformation would bend the line $y = -\pi$ at the point $x = a$ instead of at $x = 0$.

Let us now investigate the transformation effected by a function corresponding to a ζ having as factors the ζ' and ζ'' with the z changed as indicated above. We should then have

$$\frac{dZ}{dz} = \zeta = \frac{\pi}{4}\left[e^{(z+a)/4} - ie^{-(z+a)/4}\right]\left[e^{(z-a)/4} + ie^{-(z-a)/4}\right],$$

or

$$\frac{dZ}{dz} = \zeta = \frac{\pi}{4}\left[e^{z/2} + e^{-z/2}\right] + \frac{\pi i}{4}\left[e^{a/2} - e^{-a/2}\right].$$

Integrating, we have

$$Z = \frac{\pi}{2}\left[e^{z/2} - e^{-z/2}\right] + \frac{\pi i K}{4} z, \tag{3}$$

where $K = e^{a/2} - e^{-a/2} = 2 \sinh \frac{a}{2}$.

Separating the real and imaginary parts, we have

$$X = \frac{\pi}{2}\left[e^{x/2} - e^{-x/2}\right]\cos\frac{y}{2} - \frac{\pi K}{4} y, \tag{4}$$

$$Y = \frac{\pi}{2}\left[e^{x/2} + e^{-x/2}\right]\sin\frac{y}{2} + \frac{\pi K}{4} x. \tag{5}$$

When $y = \pi$,

$$X(\pi) = -\frac{\pi^2 K}{4} \quad \text{and} \quad Y(\pi) = \frac{\pi}{2}\left[e^{x/2} + e^{-x/2}\right] + \frac{\pi K}{4} x.$$

When $y = -\pi$,

$$X(-\pi) = \frac{\pi^2 K}{4} \quad \text{and} \quad Y(-\pi) = -\frac{\pi}{2}\left[e^{x/2} + e^{-x/2}\right] + \frac{\pi K}{4} x.$$

Here $Y(\pi)$ has its minimum value when $x = -a$; and $Y(-\pi)$ has its maximum value when $x = a$; that is, the upper edge of the strip on the z-plane is bent at the point $x = -a$ and the lower edge at the point $x = a$, each becoming a vertical line in the new plane extending to infinity in one direction. The perpendicular distance between these walls* is $\pi^2 K/2$. Introducing a multiplier $4/(\pi K)$ in the right-hand side of equation (3) will keep the distance between the walls a constant equal to 2π. The over-lap of the two walls is then equal to

$$2a - 4\frac{e^a + 1}{e^a - 1},$$

which becomes zero when $a = 2.40 \cdots$. By a proper choice of the a the over-lap can be made any desired amount either positive or negative. The

* The two straight lines on the Z-plane corresponding to the lines $y = \pi$ and $y = -\pi$ on the z-plane will be termed the walls.

equation corresponding to the line $y = 0$ is

$$X = \frac{4}{K} \sinh \frac{Y}{2}.$$

The drawing for the case $a = 4$ is given in figure 1, and for the case $a = 1$ in figure 2.

If we differentiate the Helmholtz equation, we obtain

$$\frac{dZ}{dz} = \zeta = 1 + e^z = (1 + ie^{z/2})(1 - ie^{z/2}).$$

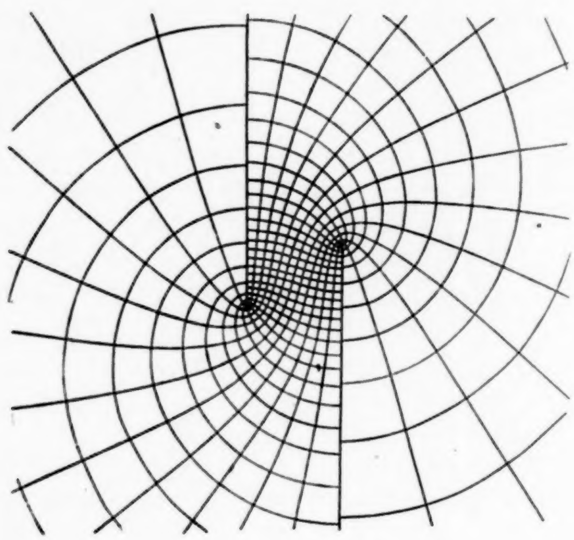


FIG. 1. $(c = \frac{1}{2},\ a = 4.)$

Now if we use the first factor only in a new ζ we have

$$\frac{dZ'}{dz} = \zeta' = (1 + ie^{z/2}),$$

$$Z' = z + 2ie^{z/2}. \tag{6}$$

Separating the real and imaginary parts, we have

$$X' = x - 2e^{x/2} \sin \frac{y}{2},$$

$$Y' = y + 2e^{x/2} \cos \frac{y}{2}.$$

When $y = \pi$,

$$X'(\pi) = x - 2e^{x/2},$$
$$Y'(\pi) = \pi.$$

Here $X'(\pi)$ has its maximum value when $x = 0$.

When $y = -\pi$,

$$X'(-\pi) = x + 2e^{x/2},$$
$$Y'(-\pi) = -\pi.$$

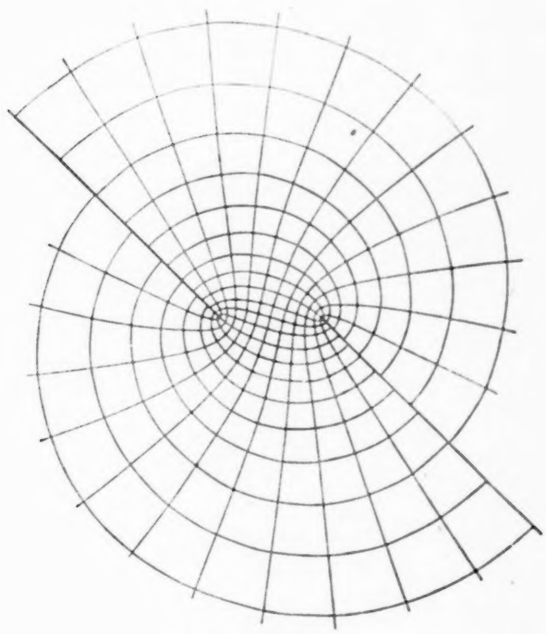


FIG. 2. ($c = \frac{1}{2}$, $a = 1$.)

Here $X'(-\pi)$ has neither a maximum nor a minimum value.

When $y = -3\pi$,

$$X'(-3\pi) = x - 2e^{x/2},$$
$$Y'(-3\pi) = -3\pi.$$

Here $X'(-3\pi)$ has its maximum value when $x = 0$. We see that here, as in the case of the transformation expressed by the hyperbolic sine, the function

derived from the first factor of ζ is similar to the original Helmholtz transformation.

If we omit the first factor of ζ we have

$$\frac{dZ''}{dz} = \zeta'' = (1 - ie^{z/2}),$$

$$Z'' = z - 2ie^{z/2}. \qquad (7)$$

Here again, as in the case of the transformation expressed by the hyperbolic sine, the Z' and the Z'' functions may be considered as each performing one half of the transformation represented by the original function from which they were obtained.

If the z in ζ' were replaced by $z + a$, equation (4) would require that the line $y = \pi$ be bent at the point $x = -a$ instead of at the point $x = 0$. If in ζ'' the z were replaced by $z - a$ equation (7) would require that the line $y = -\pi$ be bent at the point $x = a$ instead of at the point $x = 0$. Let us now use ζ' and ζ'', with the z in the former changed to $z + a$, and the z in the latter changed to $z - a$, as factors of ζ. We then have

$$\frac{dZ}{dz} = \zeta = \left[1 + ie^{(z+a)/2}\right]\left[1 - ie^{(z-a)/2}\right] = 1 + e^z + iKe^{z/2};$$

$$Z = z + e^z + 2iKe^{z/2}. \qquad (8)$$

Separating the real and imaginary parts and substituting π and $-\pi$ for x, we get

$$X(\pi) = x - e^x - 2Ke^{x/2},$$

$$Y(\pi) = \pi,$$

$$X(-\pi) = x - e^x + 2Ke^{x/2},$$

and

$$Y(-\pi) = -\pi.$$

Here $X(\pi)$ has its maximum value when $x = -a$, and $X(-\pi)$ has its maximum value when $x = a$. The upper and lower edges of the strip on the z-plane are bent at the points $x = -a$ and $x = +a$, respectively, and are transformed into two horizontal walls extending to infinity in the same direction. One wall protrudes $2a - e^{-a} + e^a$ units beyond the other with a con-

stant perpendicular distance, 2π, between them. By a proper choice of the a the protrusion can be made any desired amount. In figure 3 is given the drawing for $a = 1$.

By differentiating the Harris equation*

$$Z = \frac{\pi}{\sin \epsilon\pi}\left[\epsilon e^{(1-\epsilon)z} - (1-\epsilon)e^{-\epsilon z}\right], \tag{9}$$

where

$$0 \leqq \epsilon \leqq \tfrac{1}{2},$$

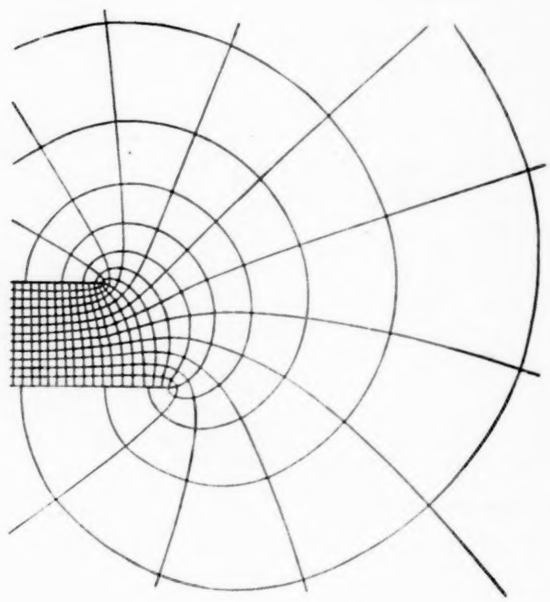


FIG. 3. ($\epsilon = 0$, $a = 1$.)

we obtain

$$\frac{dZ}{dz} = \frac{\pi\epsilon(1-\epsilon)}{\sin \epsilon\pi}\left[e^{(1-\epsilon)z} + e^{-\epsilon z}\right]$$

$$= \frac{\pi\epsilon(1-\epsilon)}{\sin \epsilon\pi}\left[e^{-\epsilon z/2} + ie^{(1-\epsilon)z/2}\right]\left[e^{-\epsilon z/2} - ie^{(1-\epsilon)z/2}\right].$$

* Harris's equation reduces to this form by the substitutions $Z = \frac{Z'}{\pi}$, $z = z' + \log \frac{\epsilon}{1-\epsilon}$.

Omitting the last factor we have

$$\frac{dZ'}{dz} = \zeta' = \frac{\pi\epsilon(1-\epsilon)}{\sin \pi\epsilon}\left[e^{-\epsilon z/2} + ie^{(1-\epsilon)z/2}\right],$$

$$Z' = \frac{\pi}{\sin \epsilon\pi}\left[i\epsilon e^{(1-\epsilon)z/2} - (1-\epsilon)\,e^{-\epsilon z/2}\right].$$

By omitting the first factor of ζ we have

$$\frac{dZ''}{dz} = \zeta'' = \frac{\pi\epsilon(1-\epsilon)}{\sin \epsilon\pi}\left[-\,ie^{(1-\epsilon)z/2} + e^{-\epsilon z/2}\right],$$

$$Z'' = \frac{-\,2\pi}{\sin \epsilon\pi}\left[i\epsilon e^{(1-\epsilon)z/2} + (1-\epsilon)e^{-\epsilon z/2}\right].$$

These functions, Z' and Z'' effect transformations similar to the Z''s and Z'''s obtained from the hyperbolic sine and the Helmholtz function. Changing the z in ζ' to $z + a$, and in ζ'' to $z - a$, and multiplying, we get

$$\frac{dZ}{dz} = \frac{\pi\epsilon(1-\epsilon)}{\sin \epsilon\pi}\left[e^{-\epsilon(z+a)/2} + ie^{(1-\epsilon)(z+a)/2}\right]\left[e^{-\epsilon(z-a)/a} - ie^{(1-\epsilon)(z-a)/2}\right]$$

$$= \frac{\pi\epsilon(1-\epsilon)}{\sin \epsilon\pi}\left[e^{(1-\epsilon)z} + e^{-\epsilon z} + iKe^{(1-2\epsilon)z/2}\right],$$

$$Z = \frac{\pi}{\sin \epsilon\pi}\left[\epsilon e^{(1-\epsilon)z} - (1-\epsilon)e^{-\epsilon z} + \frac{2iK\epsilon(1-\epsilon)}{1-2\epsilon}\,e^{(1-2\epsilon)z/2}\right] + M + iN, \quad (10)$$

where $$M = (1-2\epsilon)\left[\frac{\pi}{\sin \epsilon\pi} + 2 - \frac{K^2}{2}\right],$$

and $$N = 2\,K\left[(1-2\epsilon) - \frac{\pi\epsilon(1-\epsilon)}{(1-2\epsilon)\sin \epsilon\pi}\right].$$

This is the general equation of transformation. Choosing the additive constant, $M + iN$, in the precise form selected above, gives symmetry with respect to the origin in the new plane and keeps the walls within the finite part of the plane in the limiting cases when $\epsilon = 1/2$ and $\epsilon = 0$.

When $a = 0$ we have as a special case the Harris transformation; when $\epsilon = 0$ we have equation (8) of which the Helmholtz equation is a special case when $a = 0$; when $\epsilon = 1/2$ we have equation (3) of which the hyperbolic sine is a special case when $a = 0$.

Separating the real and imaginary parts of equation (10) we have

$$Y = \frac{\pi}{\sin \epsilon\pi}\Big[\epsilon e^{(1-\epsilon)x}\sin(1-\epsilon)y + (1-\epsilon)e^{-\epsilon x}\sin\epsilon y + \frac{2K\epsilon(1-\epsilon)}{1-2\epsilon}e^{(1-2\epsilon)x/2}\cos\frac{1-2\epsilon}{2}y\Big] + N, \tag{11}$$

$$X = \frac{\pi}{\sin \epsilon\pi}\Big[\epsilon e^{(1-\epsilon)x}\cos(1-\epsilon)y - (1-\epsilon)e^{-\epsilon x}\cos\epsilon y - \frac{2K\epsilon(1-\epsilon)}{1-2\epsilon}e^{(1-2\epsilon)x/2}\sin\frac{1-2\epsilon}{2}y\Big] + M. \tag{12}$$

When $y = \pi$,

$$X(\pi) = \frac{-\pi}{\tan \epsilon\pi}\Big[\epsilon e^{(1-\epsilon)x} + (1-\epsilon)e^{-\epsilon x} + \frac{2K\epsilon(1-\epsilon)}{1-2\epsilon}e^{(1-2\epsilon)x/2}\Big] + M, \tag{13}$$

$$Y(\pi) = \pi\Big[\epsilon e^{(1-\epsilon)x} + (1-\epsilon)e^{-\epsilon x} + \frac{2K\epsilon(1-\epsilon)}{1-2\epsilon}e^{(1-2\epsilon)x/2}\Big] + N. \tag{14}$$

Eliminating x between equations (13) and (14) we have

$$Y(\pi) = -X(\pi)\tan\epsilon\pi + M\tan\epsilon\pi + N, \tag{15}$$

the equation of a straight line in the Z-plane.

Here $X(\pi)$ has its maximum when $x = -a$. When $x = -a$, let $X(\pi)$ and $Y(\pi)$ be represented by $X(\pi, -a)$ and $Y(\pi, -a)$ respectively.

$$X(\pi, -a) = \frac{-\pi e^{\epsilon a}}{\tan\epsilon\pi}\Big[\frac{1-\epsilon-\epsilon e^{-a}}{1-2\epsilon}\Big] + M, \tag{16}$$

$$Y(\pi, -a) = \pi e^{\epsilon a}\Big[\frac{1-\epsilon-\epsilon e^{-a}}{1-2\epsilon}\Big] + N. \tag{17}$$

When $y = -\pi$,

$$X(-\pi) = \frac{-\pi}{\tan \epsilon\pi}\Big[\epsilon e^{(1-\epsilon)x} + (1-\epsilon)e^{-\epsilon x} - \frac{2K\epsilon(1-\epsilon)}{1-2\epsilon}e^{(1-2\epsilon)x/2}\Big] + M, \tag{18}$$

$$Y(-\pi) = -\pi\Big[\epsilon e^{(1-\epsilon)x} + (1-\epsilon)e^{-\epsilon x} - \frac{2K\epsilon(1-\epsilon)}{1-2\epsilon}e^{(1-2\epsilon)x/2}\Big] + N. \tag{19}$$

Eliminating x between equations (18) and (19) we have

$$Y(-\pi) = X(-\pi)\tan\epsilon\pi - M\tan\epsilon\pi + N, \tag{20}$$

the equation of a straight line in the Z-plane.

Here $X(-\pi)$ has a maximum when $x = a$. When $x = a$, let $X(-\pi)$ and $Y(-\pi)$ be represented by $X(-\pi, a)$ and $Y(-\pi, a)$ respectively.

$$Y(-\pi, a) = \frac{-\pi e^{-\epsilon a}}{\tan \epsilon\pi}\left[\frac{1 - \epsilon - \epsilon e^{a}}{1 - 2\epsilon}\right] + M, \tag{21}$$

$$Y(-\pi, a) = -\pi e^{-\epsilon a}\left[\frac{1 - \epsilon - \epsilon e^{a}}{1 - 2\epsilon}\right] + N. \tag{22}$$

From equations (16), (17), (21), and (22) we find the length of the perpendicular let fall from the end of one wall upon the other to be

$$A = \frac{2\pi e^{\epsilon a}}{1 - 2\epsilon}\left[1 - e - \epsilon e^{-a}\right]\cos \epsilon\pi,$$

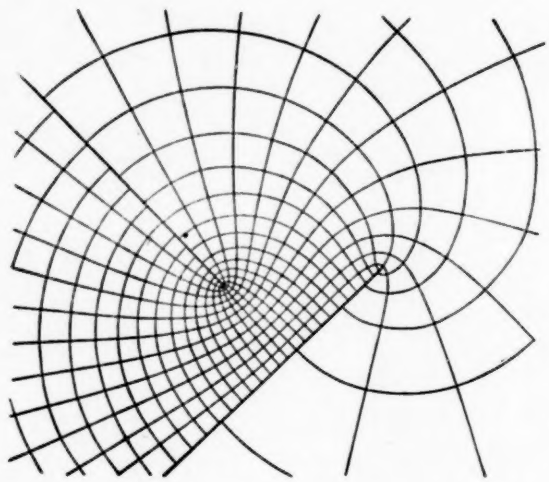


FIG. 4. ($\epsilon = \frac{1}{4}$, $a = 4$.)

and the protrusion of the second wall beyond the foot of this perpendicular to be

$$B = \frac{\pi}{(1 - 2\epsilon)\sin \epsilon\pi}\left[e^{\epsilon a}(1 - \epsilon - \epsilon e^{-a})\cos 2\epsilon\pi - e^{-\epsilon a}(1 - \epsilon - \epsilon e^{a})\right].$$

Then
$$\frac{B}{A} = \cot 2\epsilon\pi - \frac{e^{-2\epsilon a}}{\sin 2\epsilon\pi}\left[\frac{1 - \epsilon - \epsilon e^{a}}{1 - \epsilon - \epsilon e^{-a}}\right].$$

By properly choosing a this ratio may be made any desired amount for any value of ϵ. Thus we see that equation (10) represents conformally upon the entire plane a strip bounded by two parallel straight lines extending to infinity in both directions. The boundaries of the strip become any two non-intersecting right lines extending each to infinity in one direction.

Figure 4 represents the case $\epsilon = 1/4$ and $a = 4$.

In hydrodynamics this general transformation gives the continuous two-dimensional flow of a liquid in the Z-plane, having as boundaries two straight

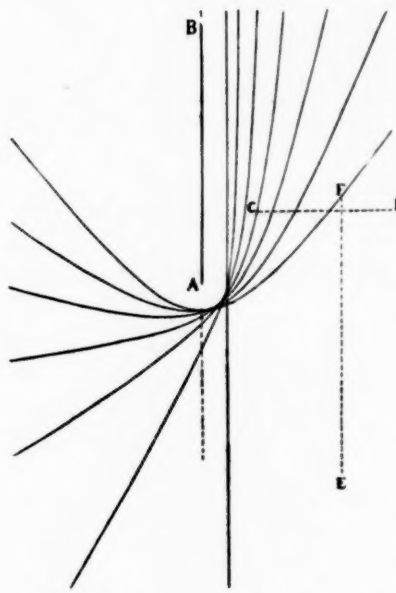


FIG. 5.

lines extending each to infinity in one direction. If the lines y = constant are taken as stream lines within the infinite strip of width 2π on the z-plane, the corresponding curves on the new plane represent stream lines passing through the opening between the two walls. Similarly, if, in the z-plane, the line $y = \pi$ is considered a source, and the line $y = -\pi$ a sink, the lines x = constant are transformed into stream lines passing from a rectilinear source to a rectilinear sink.

In figure 5, AB represents one wall held stationary while the second wall

is allowed to move about in the plane as a and ϵ vary. CD is a particular position of the second wall when at right angles to the first, i. e., $\epsilon = 1/4$. As a varies, or as the wall CD moves parallel to itself, the point C will describe a curve. Similar curves are drawn also for $\epsilon = 0, 1/12, 1/6, 1/3$, and $1/2$.

MADISON, WISCONSIN.

A PROBLEM IN CHANCE

By James K. Whittemore

A problem in the elementary theory of probabilities, the result of which will interest most students of "College Algebra," is the following:

Two players of unequal skill, A and B, play tennis. To which of them is it of advantage that deuce sets be played?

The discussion is based on the two fundamental principles for finding the probabilities of compound events: *

I. If the chances of the occurrence of two independent events are p_1 and p_2 respectively, the chance that both will occur is the product $p_1 p_2$.

II. If the chances of the occurrence of two mutually exclusive events are p_1 and p_2 respectively, the chance that one of the events will occur is the sum $p_1 + p_2$.

We shall suppose that A's chance of winning each game is fixed and equal to p. Then B's chance of winning each game is $1 - p$, and we write $1 - p = q$. This assumption is artificial but may fairly be made, and p may be chosen as the ratio of the number of games won by A to the whole number played, when the latter number is large.

We may solve the question proposed by finding A's chances of winning the set by each of the two methods of scoring, and then comparing the results.

The chance that A win the set "six-love" is, by I, equal to p^6.

The chance that he win "six-one" we find as follows: of the first six games B wins one, and A five, then A wins the seventh; the first six games may be arranged in six ways, for B may win any one of them. The chance of the occurrence of any particular one of these arrangements is p^6q. Since the different arrangements are mutually exclusive, the chance of the occurrence of one, hence that A win "six-one," is $6p^6q$.

In the same way we find that A's chances of winning with a score of "six-two," "six-three," "six-four" and "six-five" are respectively $21p^6q^2$, $56p^6q^3$, $126p^6q^4$ and $252p^6q^5$. These results are easily found if we remember that A must win the last game, that his other five games and the games won by

* See *e. g.*, Wentworth, *College Algebra*, revised edition, pp. 282, 283.

B may be arranged in any order, and that the different arrangements are mutually exclusive events.

Let us write

$$M = p^6 (1 + 6q + 21q^2 + 56q^3 + 126q^4)$$
$$Np = 252p^6q^5.$$

Then if deuce sets are not played, since winning by different scores gives again mutually exclusive events, A's chance of winning the set is

$$M + Np. \tag{1}$$

If deuce sets are to be played the result is quite different. In this case A may still win the set with B's score less than five games, and for this the chance is, as before, equal to M. But if B wins five games the score must become "five-all," for which event the chance is N. If A wins the next two games, for which the chance is p^2, he wins the set at "seven-five." The chance that he win by this score is then Np^2. If A wins the set, but not the next two games after "five-all," the score must become "six-all." The chance that A and B win each one of the two games after "five-all" is $2pq$. Then the chance that A win the set at "eight-six" is $Np^2\,2pq$. Similarly we may find the chance that A win at "nine-seven" is $Np^2\,(2pq)^2$. Hence, A's chance to win the set is

$$M + Np^2\,[1 + 2pq + (2pq)^2 + (2pq)^3 \cdots].$$

The geometric series is convergent; for $2pq = 2p(1-p) < \frac{1}{2}$, since $p \neq \frac{1}{2}$. Hence, finally, we may write A's chance of winning, when deuce sets are to be played, as

$$M + \frac{Np^2}{1 - 2pq}. \tag{2}$$

This chance is greater than his chance to win when deuce sets are not played, if

$$\frac{p^2}{1-2pq} > p,$$

that is, if $p > 1 - 2p(1-p)$, since $q = 1 - p$,

or $$(2p-1)(1-p) > 0,$$

or $$1 > p > \tfrac{1}{2}.$$

Thus the advantage of playing deuce sets lies with A if he is the better player.

As a numerical case, let A's chance of winning a single game be $p = 0.63$. Then A's chance of winning a particular set in a tournament in which deuce sets are not to be played is 0.815, while his chance of winning a particular set in a tournament in which deuce sets are to be played is 0.835. This example is chosen so that the difference between the chances in the two kinds of tournaments shall be a maximum.*

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.,
MAY, 1907.

---

* Thus, the difference in question is

$$D = 252p^6q^5\left[\frac{p}{1-2pq} - 1\right] = \frac{63}{512}\,\frac{x(1-x^2)^6}{1+x^2},$$

where $x = p - q$, and the value of x which makes D a maximum is found to be a root of the equation $11x^4 + 14x^2 - 1 = 0$, namely $x = 0.26+$. Then

$$p = \frac{1}{2} + \frac{x}{2} = 0.63+ \quad \text{and} \quad q = \frac{1}{2} + \frac{x}{2} = 0.37-,$$

and the corresponding maximum value of D is $0.02-$.

THE EXPRESSION OF CONSTANT AND OF ALTERNATING CONTINUED FRACTIONS IN HYPERBOLIC FUNCTIONS

By A. E. Kennelly

In the theory of electric telegraphic or of electric telephonic lines, it becomes necessary to consider successive fractions of the forms

$$\frac{1}{a},\quad \cfrac{1}{b+\cfrac{1}{a}},\quad \cfrac{1}{a+\cfrac{1}{b+\cfrac{1}{a}}},\quad \cfrac{1}{b+\cfrac{1}{a+\cfrac{1}{b+\cfrac{1}{a}}}}\quad \text{etc.}$$

It is desirable to obtain a simpler expression of these forms, when they are extended to a large number of terms.

Alternating continued fractions. Let us then consider two infinite continued fractions:

$$F_\infty(a,b) = \cfrac{1}{a+\cfrac{1}{b+\cfrac{1}{a+\cfrac{1}{b+\cdots}}}} \qquad F_\infty(b,a) = \cfrac{1}{b+\cfrac{1}{a+\cfrac{1}{b+\cfrac{1}{a+\cdots}}}}$$

These may be called *alternating continued fractions*, because the two quantities a and b recur alternately in the successive denominators. Denoting the successive convergents of these fractions by

$$F_1(a,b),\quad F_2(a,b),\quad F_3(a,b),\quad F_4(a,b),\ \text{etc.,} \qquad \text{and}$$

$$F_1(b,a),\quad F_2(b,a),\quad F_3(b,a),\quad F_4(b,a),\ \text{etc.,}$$

the successive values of the physical problem are

$$F_1(a,b),\quad F_2(b,a),\quad F_3(a,b),\quad F_4(b,a),\ \text{etc.}$$

Assuming first that a and b are real numbers of the same sign, we introduce two auxiliary variables θ_1 and θ_2 defined by the equations:

$$\sinh\theta_1 = \frac{a+b}{4} + \frac{a-b}{4}\sqrt{1+\frac{4}{ab}}, \quad \cosh\theta_1 = \frac{a-b}{4} + \frac{a+b}{4}\sqrt{1+\frac{4}{ab}}, \tag{1}$$

$$\sinh\theta_2 = \frac{a+b}{4} - \frac{a-b}{4}\sqrt{1+\frac{4}{ab}}, \quad \cosh\theta_2 = \frac{b-a}{4} + \frac{a+b}{4}\sqrt{1+\frac{4}{ab}}, \tag{2}$$

so that $\quad \sinh\theta_1 + \sinh\theta_2 = \dfrac{a+b}{2}; \quad \cosh\theta_1 - \cosh\theta_2 = \dfrac{a-b}{2}.$

We also define μ and δ by the equations:

$$\mu = \frac{\theta_1+\theta_2}{2} = \sinh^{-1}\frac{\sqrt{ab}}{2} = \cosh^{-1}\sqrt{1+\frac{ab}{4}}, \tag{3}$$

$$\delta = \frac{\theta_1-\theta_2}{2} = \sinh^{-1}\left(\frac{a-b}{2\sqrt{ab}}\right) = \cosh^{-1}\left(\frac{a+b}{2\sqrt{ab}}\right). \tag{4}$$

In these equations the radicals are understood to have a positive sign, and consequently μ is always positive.

We now have:

$$F_1(a, b) = \frac{1}{a} = \frac{1}{\sqrt{ab}}\cdot\sqrt{\frac{b}{a}} = \frac{1}{2\sinh\mu}\sqrt{\frac{b}{a}}$$

$$= \frac{\cosh\mu}{\sinh 2\mu}\sqrt{\frac{b}{a}} = \frac{\cosh\mu}{\sinh 2\mu}(\cosh\delta - \sinh\delta),$$

$$F_2(a, b) = \frac{\sinh 2\mu}{\cosh 3\mu}\sqrt{\frac{b}{a}} = \frac{\sinh 2\mu}{\cosh 3\mu}(\cosh\delta - \sinh\delta),$$

$$F_3(a, b) = \frac{\cosh 3\mu}{\sinh 4\mu}\sqrt{\frac{b}{a}} = \frac{\cosh 3\mu}{\sinh 4\mu}(\cosh\delta - \sinh\delta).$$

In general,

$$F_n(a, b) = \frac{\sinh n\mu}{\cosh(n+1)\mu}(\cosh\delta - \sinh\delta) = \frac{\sinh n\mu}{\cosh(n+1)\mu}\sqrt{\frac{b}{a}} \text{ if } n \text{ even}; \tag{5}$$

$$F_n(a, b) = \frac{\cosh n\mu}{\sinh(n+1)\mu}(\cosh\delta - \sinh\delta) = \frac{\cosh n\mu}{\sinh(n+1)\mu}\sqrt{\frac{b}{a}} \text{ if } n \text{ odd}. \tag{6}$$

These formulas may be proved by induction. Assuming the truth of the formula for $F_{n-1}(a, b)$, and the corresponding formula* for $F_{n-1}(b, a)$ in which a and b exchange places, we have, if n is odd,

$$F_n(a, b) = \frac{1}{a + F_{n-1}(b, a)} = \frac{1}{a + \sqrt{\frac{a}{b}} \cdot \frac{\sinh (n-1)\mu}{\cosh n\mu}}$$

$$= \frac{1}{\left\{\sqrt{ab} + \frac{\cosh \mu \sinh n\mu - \sinh \mu \cosh n\mu}{\cosh n\mu}\right\}\sqrt{\frac{a}{b}}}$$

$$= \frac{1}{(2 \sinh \mu + \cosh \mu \tanh n\mu - \sinh \mu)\sqrt{\frac{a}{b}}} = \frac{1}{(\sinh \mu + \cosh \mu \tanh n\mu)\sqrt{\frac{a}{b}}}$$

$$= \frac{1}{\frac{\sinh n\mu \cosh \mu + \cosh n\mu \sinh \mu}{\cosh n\mu}\sqrt{\frac{a}{b}}} = \frac{\cosh n\mu}{\sinh \{(n+1)\mu\}}\sqrt{\frac{b}{a}};$$

and if n is even,

$$F_n(a, b) = \frac{1}{a + F_{n-1}(b, a)} = \frac{1}{a + \sqrt{\frac{a}{b}} \cdot \frac{\cosh (n-1)\mu}{\sinh n\mu}}$$

$$= \frac{1}{\left\{\sqrt{ab} + \frac{\cosh \mu \cosh n\mu - \sinh \mu \sinh n\mu}{\sinh n\mu}\right\}\sqrt{\frac{a}{b}}}$$

$$= \frac{1}{(2 \sinh \mu + \cosh \mu \coth n\mu - \sinh \mu)\sqrt{\frac{a}{b}}} = \frac{1}{(\sinh \mu + \cosh \mu \coth n\mu)\sqrt{\frac{a}{b}}}$$

$$= \frac{1}{\frac{\cosh n\mu \cosh \mu + \sinh n\mu \sinh \mu}{\sinh n\mu}\sqrt{\frac{a}{b}}} = \frac{\sinh n\mu}{\cosh \{(n+1)\mu\}}\sqrt{\frac{b}{a}}.$$

---

* It follows at once that $F_n(b, a) = \frac{\sinh n\mu}{\cosh (n+1)\mu}\sqrt{\frac{a}{b}} = \frac{\sinh n\mu}{\cosh (n+1)\mu}(\cosh \delta + \sinh \delta)$

or $F_n(b, a) = \frac{\cosh n\mu}{\sinh (n+1)\mu}\sqrt{\frac{a}{b}} = \frac{\cosh n\mu}{\sinh (n+1)\mu}(\cosh \delta + \sinh \delta)$

according as n is even or odd.

In none of these equations is either sinh $n\mu$ or cosh $n\mu$ equal to zero, since μ is a positive real number.

If we define the value of the infinite continued fraction as

$$F_\infty(a, b) = \lim_{n=\infty} F_n(a, b),$$

we have, since μ is positive,

$$F_\infty(a, b) = \epsilon^{-\theta_1} = \epsilon^{-(\mu+\delta)} = \sqrt{\frac{b}{a}} \cdot \epsilon^{-\mu} = \cosh\theta_1 - \sinh\theta_1$$

$$= \frac{b}{2}\left(\sqrt{1 + \frac{4}{ab}} - 1\right). \tag{7}$$

The last result may also be obtained by elementary **algebra**, which offers, however, no proof of the existence of the limit.

It may also be noted that $\dfrac{F_n(b, a)}{F_n(a, b)} = \dfrac{a}{b}$, whether n be odd or even.

Similarly,

$$F_\infty(b, a) = \epsilon^{-\theta_2} = \epsilon^{-(\mu-\delta)} = \sqrt{\frac{a}{b}} \cdot \epsilon^{-\mu} = \cosh\theta_2 - \sinh\theta_2$$

$$= \frac{a}{2}\left(\sqrt{1 + \frac{4}{ab}} - 1\right). \tag{8}$$

Constant continued Fractions. If $b = a$, each of the alternating continued fractions $F_\infty(a, b)$ and $F_\infty(b, a)$ becomes identical with

$$F_\infty(a) = \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a + \cdots}}}$$

which we may call a *constant* continued fraction. In this case

$$\theta_1 = \theta_2 = \mu = \theta = \sinh^{-1}\left(\frac{a}{2}\right), \text{ and } \delta = 0.$$

The successive convergents then become:

$$F_1(a) = \frac{\cosh\theta}{\sinh 2\theta},$$

$$F_2(a) = \frac{\sinh 2\theta}{\cosh 3\theta},$$

$$F_3(a) = \frac{\cosh 3\theta}{\sinh 4\theta},$$

. . . . . .

$$F_n(a) = \frac{\sinh n\theta}{\cosh (n+1)\theta} \qquad \text{if } n \text{ is even,} \tag{9}$$

$$= \frac{\cosh n\theta}{\sinh (n+1)\theta} \qquad \text{if } n \text{ is odd;} \tag{10}$$

. . . . . .

$$F_\infty(a) = \epsilon^{-\theta}$$

$$= \cosh\theta - \sinh\theta = \frac{a}{2}\left(\sqrt{1+\frac{4}{a^2}} - 1\right). \tag{11}$$

From the above expression for $F_\infty(a)$, it follows that any exponential value ϵ^{-m} may be regarded as the value of a constant continued fraction $F_\infty(2\sinh m)$.

Alternative Expression for a Constant Continued Fraction. The following alternative form for $F_n(a)$ is sometimes convenient.

$$F_1(a) = \frac{1}{\sinh\theta + \cosh\theta\tanh\theta} = \frac{1}{\sinh\theta + \sinh\theta} = \frac{1}{2\sinh\theta} = \frac{1}{a}, \tag{12}$$

$$F_2(a) = \frac{1}{\sinh\theta + \cosh\theta\coth 2\theta},$$

$$F_3(a) = \frac{1}{\sinh\theta + \cosh\theta\tanh 3\theta},$$

. . . . . .

$$F_n(a) = \frac{1}{\sinh\theta + \cosh\theta\coth n\theta} \qquad \text{if } n \text{ is even,}$$

$$= \frac{1}{\sinh\theta + \cosh\theta\tanh n\theta} \qquad \text{if } n \text{ is odd;}$$

. . . . . .

$$F_\infty(a) = \frac{1}{\epsilon^\theta} = \epsilon^{-\theta} = \epsilon^{-\sinh^{-1}\left(\frac{a}{2}\right)} = \frac{1}{\sinh\theta + \cosh\theta} = \cosh\theta - \sinh\theta \tag{13}$$

Consequently, as n increases from 1 to ∞, the second term in the denominator of (12) changes from $\sinh\theta$ to $\cosh\theta$ in (13).

Expression of any alternating continued fraction in terms of an equivalent constant continued fraction. It is evident on comparing equations (5) and (6) with (9) and (10) that the nth convergent of either of the alternating continued fractions $F(a, b)$ and $F(b, a)$ may be expressed in terms of the nth convergent of the constant continued fraction $F(\sqrt{ab})$, the constant term of which is the geometric mean of the two alternating terms in the alternating continued fraction. Thus:

$$F_5(a, b) = \cfrac{1}{a + \cfrac{1}{b + \cfrac{1}{a + \cfrac{1}{b + \cfrac{1}{a}}}}} = \sqrt{\frac{b}{a}} \cdot F_5(\sqrt{ab}) = \frac{b}{\sqrt{ab}} \times \cfrac{1}{\sqrt{ab} + \cfrac{1}{\sqrt{ab} + \cfrac{1}{\sqrt{ab} + \cfrac{1}{\sqrt{ab} + \cfrac{1}{\sqrt{ab}}}}}}$$

and

$$F_5(b, a) = \cfrac{1}{b + \cfrac{1}{a + \cfrac{1}{b + \cfrac{1}{a + \cfrac{1}{b}}}}} = \sqrt{\frac{a}{b}} \cdot F_5(\sqrt{ab}) = \frac{a}{\sqrt{ab}} \times \cfrac{1}{\sqrt{ab} + \cfrac{1}{\sqrt{ab} + \cfrac{1}{\sqrt{ab} + \cfrac{1}{\sqrt{ab} + \cfrac{1}{\sqrt{ab}}}}}}$$

Particular cases of Alternating Continued Fractions. In the case of $b = \dfrac{a}{a+1}$ in $F_\infty(b, a)$, we find by (2) that $\theta_2 = 0$ and $F_\infty(b, a) = 1$.

Similarly, when $a = \dfrac{b}{b+1}$, we find $\theta_1 = 0$ and $F_\infty(a, b) = 1$.

Case of a and b with opposite signs. If a and b have opposite signs, let $\sqrt{ab} = ic$, where c is a positive real number and $i = \sqrt{-1}$. Defining μ as in (3), we have

$$\sinh\mu = \frac{ic}{2},$$

whence,

$$\epsilon^{\mu} - \epsilon^{-\mu} = ic.$$

It follows, if we write $\mu = \nu + i\lambda$, that either $\nu = 0$, or $\lambda = \frac{\pi}{2}$ and ν is positive. If $\nu = 0$ we may show that

$$ab = -4\sin^2\lambda,$$

and hence that ab lies between 0 and -4. In this case the successive values of the convergents approach no limit, as may be readily verified. For all other cases

$$\mu = \nu + i\frac{\pi}{2}.$$

Substituting in (3) we obtain

$$\nu = \cosh^{-1}\left(\frac{c}{2}\right) = \sinh^{-1}\sqrt{\frac{c^2}{4} - 1}\,, \tag{14}$$

where again the radical has the positive sign, since ν is positive. Similarly, from (4),

$$\delta = \gamma \pm i\frac{\pi}{2}, \tag{15}$$

where γ is a real quantity, the sign of $i\frac{\pi}{2}$ following the sign of $(a - b)$ in (4). Substituting these values of μ and δ in equations (5) and (6) we find, whether n be odd or even:

$$F_n(a, b) = -i\frac{\sinh n\nu}{\sinh(n+1)\nu}\sqrt{\frac{b}{a}}\,, \tag{16}$$

and

$$F_\infty(a, b) = -i\sqrt{\frac{b}{a}}\;.\;\epsilon^{-\nu} = \frac{b}{2}\left(\sqrt{1 + \frac{4}{ab}} - 1\right). \tag{17}$$

It also follows that, whether n be odd or even,

$$F_n(b, a) = -i \frac{\sinh n\nu}{\sinh (n+1)\nu} \sqrt{\frac{a}{b}}, \tag{18}$$

and
$$F_\infty(b, a) = -i\sqrt{\frac{a}{b}} \cdot \epsilon^{-\nu} = \frac{a}{2}\left(\sqrt{1 + \frac{4}{ab}} - 1\right). \tag{19}$$

Case of $b = -a$, *and* $-ab = c^2 > 4$. If $b = -a$, $F(a, b)$ becomes

$$F(a, b) = \cfrac{1}{a + \cfrac{1}{-a + \cfrac{1}{a + \cfrac{1}{-a + \cdots}}}} = \cfrac{1}{a - \cfrac{1}{a - \cfrac{1}{a - \cfrac{1}{a - \cdots}}}}$$

so that a constant continued fraction with all its signs negative is equivalent to an alternating continued fraction with its two terms equal in magnitude but opposite in sign. In this case, whether n be odd or even,

$$F_n(a, -a) = \frac{\sinh n\nu}{\sinh (n+1)\nu}, \tag{20}$$

and
$$F_\infty(a, -a) = \epsilon^{-\nu} = -\frac{a}{2}\left(\sqrt{1 - \frac{4}{a^2}} - 1\right). \tag{21}$$

Since $F_n(-a, a) = -F_n(a, -a)$, the successive convergents of $F_n(-a, a)$ are found by taking the results in (20) and (21) with the opposite sign. Moreover, when $b = -a$, $F(\sqrt{ab})$ becomes $F(ic)$ or

$$F(ic) = \cfrac{1}{ic + \cfrac{1}{ic + \cfrac{1}{ic + \cfrac{1}{ic + \cdots}}}} = -i \times \cfrac{1}{c - \cfrac{1}{c - \cfrac{1}{c - \cfrac{1}{c - \cdots}}}} = -i \times \cfrac{1}{c + \cfrac{1}{-c + \cfrac{1}{c + \cfrac{1}{-c + \cdots}}}}$$

so that substituting in (20) and (21), whether n be odd or even,

$$F_n(ic) = -i \frac{\sinh n\nu}{\sinh (n+1)\nu} \tag{22}$$

and
$$F_\infty(ic) = -i\epsilon^{-\nu} = \frac{ic}{2}\left(\sqrt{1-\frac{4}{c^2}}-1\right). \tag{23}$$

Case of a and b having opposite signs, with $0 < c^2 < 4$. In this case we may take $\nu = 0$, and write $\mu = i\lambda$, where λ is a positive real number. Then equations (16) and (17) become:

$$F_n(a,b) = -i\sqrt{\frac{b}{a}} \times \frac{\sin n\lambda}{\sin (n+1)\lambda}, \tag{24}$$

$$F_n(b,a) = -i\sqrt{\frac{a}{b}} \times \frac{\sin n\lambda}{\sin (n+1)\lambda}, \tag{25}$$

where $\lambda = \cos^{-1}\left(\frac{c}{2}\right) = \sin^{-1}\sqrt{1-\frac{c^2}{4}}$, whether n be odd or even. As already pointed out, these expressions do not converge to a limit.

Case of $b = -a$ with $0 < c^2 < 4$.

When $b = -a$ in (24) we have

$$F_n(a, -a) = \frac{\sin n\lambda}{\sin (n+1)\lambda} = \cfrac{1}{a - \cfrac{1}{a - \cfrac{1}{a - \cdots}}} \text{ to } n \text{ terms.} \tag{26}$$

and
$$F_n(-a, a) = \frac{-\sin n\lambda}{\sin (n+1)\lambda}, \tag{27}$$

whether n be odd or even. These expressions do not converge as n becomes infinite.

Equation (26) covering the particular case of a constant continued fraction with all its numerators negative after the first, and $0 < c^2 < 4$, expressed in terms of circular functions, was published by Strehlke in 1864.* Strehlke's formula does not, however, extend to $F_n(a)$, a constant continued fraction with positive signs, and still less does it apply to alternating continued fractions.

* A. B. Strehlke, Grunert's *Archiv der Mathematik und Physik*, 1864, vol. 42, p. 343.

Terminally Loaded Alternating Continued Fractions. Among the physical applications already referred to, the following *ascending* series of alternating continued fractions present themselves:

$$F_1\,(a,\,b)_m = \frac{1}{a+m}$$

$$F_2(b,\,a)_m = \frac{1}{b+\dfrac{1}{a+m}}$$

$$F_3(a,\,b)_m = \frac{1}{a+\dfrac{1}{b+\dfrac{1}{a+m}}} \quad \text{etc.}$$

Consistently with this notation, we may write

$$F_0(b,\,a)_m = m.$$

The constant quantity m appearing at the end of each expression in this series may be called the *terminal load* of the alternate continued fractions $F_n(a,\,b)_m$ and $F_n(b,\,a)_m$.

It may be demonstrated that:

$$F_n(a,\,b)_m = \sqrt{\frac{b}{a}}\cdot\frac{\sinh\,(n\mu+\phi)}{\cosh\,\{(n+1)\,\mu+\phi\}} \qquad \text{if } n \text{ is even} \tag{28}$$

$$= \sqrt{\frac{b}{a}}\cdot\frac{\cosh\,(n\mu+\phi)}{\sinh\,\{(n+1)\,\mu+\phi\}} \qquad \cdot\cdot\cdot \text{ odd} \tag{29}$$

$$F_n(b,\,a)_m = \sqrt{\frac{a}{b}}\cdot\frac{\sinh\,(n\mu+\phi)}{\cosh\,\{(n+1)\,\mu+\phi\}} \qquad \cdot\cdot\cdot \text{ even} \tag{30}$$

$$= \sqrt{\frac{a}{b}}\cdot\frac{\cosh\,(n\mu+\phi)}{\sinh\,\{(n+1)\,\mu+\phi\}} \qquad \cdot\cdot\cdot \text{ odd} \tag{31}$$

where μ is the same quantity as is defined in (3) or appears in equations (5) and (6) for the corresponding unloaded alternating continued fractions, and

ϕ is defined by the relation :

$$F_0(b,\ a)_m = m = \sqrt{\frac{a}{b}} \cdot \frac{\sinh\phi}{\cosh(\mu+\phi)}$$

or

$$\phi = \tanh^{-1} \frac{\cosh\mu}{\sqrt{\frac{a}{b}} \cdot \frac{1}{m} - \sinh\mu}. \tag{32}$$

If m be a positive real quantity between 0 and $\sqrt{\frac{a}{b}} \cdot \epsilon^{-\mu} = F_\infty(b,\ a)$, ϕ is real and positive. If m exceeds $\sqrt{\frac{a}{b}} \cdot \epsilon^{-\mu}$, we may write

$$\phi = \phi' + \frac{i\pi}{2} \tag{33}$$

or

$$\phi' = \tanh^{-1} \left\{ \frac{\sqrt{\frac{a}{b}} \cdot \frac{1}{m} - \sinh\mu}{\cosh\mu} \right\}, \tag{34}$$

where ϕ' is positive and real.

Substituting in equations (28) to (31), we obtain when m exceeds $\sqrt{\frac{a}{b}} \cdot \epsilon^{-\mu}$,

$$F_n(a,\ b)_m = \sqrt{\frac{b}{a}} \cdot \frac{\cosh(n\mu+\phi')}{\sinh\{(n+1)\mu+\phi'\}} \qquad \text{if } n \text{ is even} \tag{35}$$

$$= \sqrt{\frac{b}{a}} \cdot \frac{\sinh(n\mu+\phi')}{\cosh\{(n+1)\mu+\phi'\}} \qquad \text{. . . odd} \tag{36}$$

$$F_n(b,\ a)_m = \sqrt{\frac{a}{b}} \cdot \frac{\cosh(n\mu+\phi')}{\sinh\{(n+1)\mu+\phi'\}} \qquad \text{. . . even} \tag{37}$$

$$= \sqrt{\frac{a}{b}} \cdot \frac{\sinh(n\mu+\phi')}{\cosh\{(n+1)\mu+\phi'\}} \qquad \text{. . . odd} \tag{38}$$

In the particular case when $m = \frac{2}{b}$, $\phi' = 0$ and

$$F_n(a, b)_{\frac{2}{b}} = \sqrt{\frac{b}{a}} \cdot \frac{\cosh n\mu}{\sinh(n+1)\mu} \qquad \text{if } n \text{ is even} \tag{39}$$

$$= \sqrt{\frac{b}{a}} \cdot \frac{\sinh n\mu}{\cosh(n+1)\mu} \qquad \text{. . . odd} \tag{40}$$

$$F_n(b, a)_{\frac{2}{b}} = \sqrt{\frac{a}{b}} \cdot \frac{\cosh n\mu}{\sinh(n+1)\mu} \qquad \text{. . . even} \tag{41}$$

$$= \sqrt{\frac{a}{b}} \cdot \frac{\sinh n\mu}{\cosh(n+1)\mu} \qquad \text{. . . odd} \tag{42}$$

The corresponding expressions for terminally loaded constant continued fractions may be found from the preceding equations by putting $b = a$.

It is evident from the above conditions that continued fractions which are not constant may be expressed in hyperbolic functions. Thus the five-stage continued fraction

$$\cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{b + \cfrac{1}{b}}}}} = F_3(a)_{F_2(b)}$$

may be expressed as the third approximation to a constant continued fraction, terminally loaded by the second approximation to another constant continued fraction.

Harvard University,
Cambridge, Mass.

A NECESSARY CONDITION THAT ALL THE ROOTS OF AN ALGEBRAIC EQUATION BE REAL

By Oliver D. Kellogg

In the Annals of Mathematics for July 1903, Professor Van Vleck has given a sufficient condition that an algebraic equation with real coefficients

$$c_0 + c_1x + c_2x^2 + \cdots + c_nx^n = 0 \tag{1}$$

have the maximum number of imaginary roots, that is, n or $n - 1$ according as n is even or odd. While the condition is not so stated, it appears from the proof that it is sufficient if the coefficients of even index, c_{2i}, and the determinants

$$\begin{vmatrix} c_{2i}, & c_{2i+1} \\ c_{2i+1}, & c_{2i+2} \end{vmatrix},$$

are all positive. It is interesting to remark that if any one of the determinants,

$$\begin{vmatrix} c_{j-1}, & c_j \\ c_j, & c_{j+1} \end{vmatrix},$$

among which the above are found, is positive, or if any vanishes for $k < j < n$, where c_k is the first coefficient of (1) different from zero, then the equation is sure to have *some* imaginary roots. This fact is merely another aspect of the following theorem:

If all the roots of the equation (1) *are real, all the determinants*

$$D_j = \begin{vmatrix} c_{j-1}, & c_j \\ c_j, & c_{j+1} \end{vmatrix} \qquad (k < j < n)$$

are negative.

This is a result of a well known corollary to Des Cartes' Rule of Signs, to the effect that if all the roots of an algebraic equation are real, and if any intermediate coefficient is zero, the adjacent coefficients must be different from

zero and have opposite signs. From this it follows at once that if any c_j vanishes, the corresponding $D_j = c_{j-1}\, c_{j+1} < 0$. Moreover, it is evident that $D_k = -c_k^2 < 0$, this inequality holding even for $k = 0$ if we define c_{-1} as zero.

If c_j is different from zero, introduce into the equation (1) a new real root a, obtaining

$$(2) \qquad (c_0 a + c_{-1}) + (c_1 a + c_0)\, x + (c_2 a + c_1)\, x^2 + \cdots + c_n x^n = 0.$$

Then choosing a so that the coefficient $c_j a + c_{j-1} = 0$ and applying the corollary cited to the equation (2) we find

$$\left(-\frac{c_{j-1}^2}{c_j} + c_{j-2}\right)\left(-\frac{c_{j+1}c_{j-1}}{c_j} + c_j\right) < 0,$$

that is,

$$D_{j-1}\, D_j > 0,$$

so that D_j and D_{j-1} are different from zero and have the same signs. Hence all the determinants in the series have the same sign, and since D_k is negative, it follows that D_j is negative for all values of j as was to be proved.

If instead of forming equation (2) we form from (1) an equation with two new real roots, and express the condition that two successive coefficients of the new equation cannot vanish, we shall be led to the theorem:

If all the roots of the equation (1) *are real, none of the quadratics*

$$\begin{vmatrix} x^2, & x, & 1 \\ c_{j-1}, & c_j, & c_{j+1} \\ c_j, & c_{j+1}, & c_{j+2} \end{vmatrix} = 0 \qquad (k \leqq j < n, \quad c_{-1} = 0, \quad c_{n+1} = 0)$$

has real roots.

COLUMBIA, MISSOURI,
JULY, 1907.

THE EQUILIBRIUM OF A HEAVY HOMOGENEOUS CHAIN IN A UNIFORMLY ROTATING PLANE

BY EDWIN BIDWELL WILSON

1. Introduction. The problem of the equilibrium of a chain has attracted considerable attention since the days when James Bernoulli first solved the simple case of a uniform chain hanging by both ends in the constant field of gravity. A large amount of theory has been developed * and a great variety of problems has been solved. It appears, however, that for the most part the cases which have yielded integrable differential equations belong to two types: Those of equilibrium in a parallel field or in a central field.† Both these fields are rectilinear, so that the question of the equilibrium of a chain with one end free is solved intuitively. The first plane curvilinear field to occur to one would probably be that of a constant force parallel to a line combined with a force at right angles to the line and varying directly as the distance from the line. Furthermore precisely such a field is afforded by the idealization of a common physical phenomenon.

Consider a uniform chain or string suspended by one end in the field of gravity and constrained to rotate by twisting that end. It is well known that the chain will either swing entirely to one side of the vertical or will take the form of a series of arches which increase in height from the fixed toward the free end. The motion of the chain rapidly approaches a steady state in which each point is revolving at a constant speed in a horizontal circle concentric with the vertical axis through the point of support. The effects of torsion and of friction against the air will somewhat disturb the experiment and very greatly complicate the differential equations which determine the figure of the curve. If we idealize the problem by omitting these disturbing effects and by assuming the chain to be uniform and mathematically perfect, that is, of no appreciable thickness and quite flexible and inelastic, we have a

* Perhaps the best account of the theory which is readily accessible is that given by Appell in his *Traité de mécanique rationnelle*, volume I, chapter 6, pages 180–204. A shorter treatment is found in Gray's *Treatise on Physics*, volume I, chapter 8, pages 286–295.

† For numerous problems see Appell, loc. cit., or Walton's *Problems in Theoretical Mechanics*, third edition, pages 114–143.

problem easy to state in differential equations. It is this problem with some generalization which will be discussed in the following pages.

According to a fundamental principle of dynamics, the uniform rotation of a system may be done away if the corresponding centrifugal force be introduced. As the centrifugal force varies directly with the distance from the axis of rotation, the problem of the equilibrium of a rotating chain reduces to that of equilibrium in the curvilinear field above mentioned, provided only that the chain lies entirely in one plane. Whether or not this condition be fulfilled depends on the initial conditions. The equations of equilibrium are

$$\frac{d}{ds}\left(T\frac{dx}{ds}\right)+\rho X=0,\quad \frac{d}{ds}\left(T\frac{dy}{ds}\right)+\rho Y=0,\quad \frac{d}{ds}\left(T\frac{dz}{ds}\right)+\rho Z=0,$$

where T is the tension, ρ the density, and X, Y, Z the external forces per unit mass at any point of the chain.

If the x-axis be vertical and directed upwards, the equations for the present problem become

$$\frac{d}{ds}\left(T\frac{dx}{ds}\right)-\rho g=0,\quad \frac{d}{ds}\left(T\frac{dy}{ds}\right)+\rho\omega^2 y=0,\quad \frac{d}{ds}\left(T\frac{dz}{ds}\right)+\rho\omega^2 z=0,$$

where ω is the angular velocity. It will be worth while to show that if the chain is attached at one point to the axis of rotation, it must lie entirely in a vertical plane through the axis. For this purpose the cylindrical equations of the chain are useful. These may be deduced exactly as the Cartesian equations are usually deduced, i. e., from the formulas

$$d(T\alpha)+\rho F_1 ds=0,\quad d(T\beta)+\rho F_2 ds=0,\quad d(T\gamma)+\rho F_3 ds=0,$$

where F_1, F_2, F_3 represent the forces along any three mutually perpendicular directions and α, β, γ are the direction cosines of the element of arc with reference to those directions. If the directions are X and, in the yz-plane, the radial direction R and the perpendicular to R, we have

$$\frac{d}{ds}\left(T\frac{dx}{ds}\right)+\rho X=0,\quad \frac{d}{ds}\left(T\frac{dr}{ds}\right)+\rho R=0,\quad \frac{d}{ds}\left(T\frac{r\,d\theta}{ds}\right)+\rho\Theta=0,$$

where X, R, Θ now represent the components of the force. In the case in hand $X=-g$, $R=\omega^2 r$, $\Theta=0$. Hence $T\,r d\theta/ds=$ const., and if the chain has a point on the axis of rotation, the constant vanishes. Hence Θ is

constant and the chain must lie in a vertical plane through the axis. Hereafter, as stated in the title, we shall restrict the discussion to this case where the chain lies in a plane.

2. The differential equations of the problem. Let the chain lie in the xy-plane and assume that the x-axis is vertical with the positive end up. The equations of the chain then reduce to

$$\frac{d}{ds}\left(T\frac{dx}{ds}\right)-\rho g=0,\qquad \frac{d}{ds}\left(T\frac{dy}{ds}\right)+\rho\omega^2 y=0 \tag{1}$$

in rectangular coordinates. The intrinsic equations are

$$\frac{dT}{ds}+\rho S=0,\qquad T+\rho RN=0, \tag{2}$$

where S and N are the tangential and normal components of the applied forces and R the radius of curvature of the chain. In the case in hand the forces are derivable from the potential.

$$V=gx-\tfrac{1}{2}\omega^2 y^2,\qquad S=-\partial V/\partial s.$$

Hence the first equation of (2) may be integrated to

$$T-\rho V=T-\rho gx+\tfrac{1}{2}\rho\omega^2 y^2=\text{const}. \tag{3}$$

If T_0 be the tension in the chain at the origin $(0, 0)$, we have

$$T=T_0+\rho gx-\tfrac{1}{2}\rho\omega^2 y^2. \tag{3'}$$

Inasmuch as T cannot become negative, it is apparent that the only region of the plane in which the chain may lie is the interior of the parabola

$$T=T_0+\rho gx-\tfrac{1}{2}\rho\omega^2 y^2=0, \tag{4}$$

which has its vertex downwards and situated at the point $x=-T_0/\rho g$.

The second of the intrinsic equations (2) may be used to obtain the differential equation of the second order for the chain. For

$$N=g\sin\tau+\omega^2 y\cos\tau,$$

where τ is the inclination of the curve to the x-axis. Hence

$$T+\rho(g\sin\tau+\omega^2 y\cos\tau)\,\frac{(1+y'^2)^{\frac{3}{2}}}{y''}=0.$$

Or

$$y'' = -\frac{\rho\omega^2 y + \rho g y'}{T_0 + \rho g x - \frac{1}{2}\rho\omega^2 y^2}(1 + y'^2). \tag{5}$$

It should be noted that this equation contains the quantity T_0 which is unknown: for with a given chain of length l suspended either at both ends or only at one, the tension at every point is a consequence of the position of the chain and is not to be counted among the initial conditions. In fact the tension at the point of support is a function of l. The differential equation of the chain is really of the third order and may be found by eliminating T_0 from (5) and that equation differentiated with respect to x. The form containing T_0 will, however, answer our purposes.*

Another method of treating the problem is by the calculus of variations. The condition for equilibrium is that the potential energy be an extremum — a minimum if the equilibrium is to be stable. As the length of the chain is constant the problem is one in relative minima; in fact we have

$$\int V ds \quad \text{a minimum,} \qquad \int ds \quad \text{constant.}$$

Hence it is necessary to minimize the integral

$$I = \int (\lambda g + gx - \tfrac{1}{2}\omega^2 y^2)\, ds, \tag{6}$$

where λ is a parameter. By a theorem due to Weierstrass,† an integral of

---

* Equation (5) can be integrated once by introducing the arc s and its derivative ds/dx. The integral may be found more readily by operating directly on the fundamental equations (1). The result is, if the point of support of the chain be at the origin,

$$(\rho g s + T_0 \cos\tau_0)\, ds/dx = T_0 + \rho g x - \tfrac{1}{2}\rho\omega^2 y^2,$$

where τ_0 is the inclination of the string at the origin. The derivation of this result, which is equivalent to our (7) below, will be left to the reader who may thus entirely avoid the use of the calculus of variations in arriving at this important equation (7). It should be noted here again that τ_0 is not properly to be considered among the initial conditions any more than T_0; it is indeed a function of l.

† Instead of reverting to Weierstrass's treatment, we may note that integrals of the form (6) are particularly simple to treat by methods recently developed by Bliss in the *Transactions of the American Mathematical Society*, volume 8 (1907), pages 405–414. He considers the standard form of the integrand in the calculus of variations as $F(x, y, \tau)ds$. This is highly convenient for the present problem: for here F does not contain τ explicitly and moreover the extremals may become perpendicular to the x-axis which is not allowed in case the integrand has the form $F(x, y, y')dx$. Bliss obtains the necessary condition $F + F_{\tau\tau} \geqq 0$ for a minimum — a condition which in this case restricts (x, y) to the interior or contour of the parabola (4').

this form cannot be a minimum if the coefficient of ds becomes negative. This shows that the only possible region for minimizing extremals is the interior of the parabola

$$\lambda g + gx - \tfrac{1}{2}\omega^2 y^2 = 0. \tag{4'}$$

Hence it appears, as may readily be shown also by deriving the differential equation of the extremals, that λ is to be identified with $T_0/\rho g$, and the fact that T_0 is not to be regarded as among the initial conditions but to be determined from the length l of the chain is further emphasized.

Some interesting and important results may be obtained by applying the ordinary rules of variation directly to (6).* For this purpose we may indicate the limits by 0 and 1 and consider them as variable.

$$\delta I = \int_0^1 \left\{ (g\delta x - \omega^2 y \delta y) ds + (\lambda g + gx - \tfrac{1}{2}\omega^2 y^2) \frac{dx\,\delta dx + dy\,\delta dy}{ds} \right\}.$$

By integration by parts we have

$$\delta I = \int_0^1 \left\{ g ds - d\left[(\lambda g + gx - \tfrac{1}{2}\omega^2 y^2) \frac{dx}{ds} \right] \right\} \delta x$$

$$+ \int_0^1 \left\{ -\omega^2 y ds - d\left[\lambda g + gx - \tfrac{1}{2}\omega^2 y^2) \frac{dy}{ds} \right] \right\} \delta y$$

$$+ \left[(\lambda g + gx - \tfrac{1}{2}\omega^2 y^2) \frac{dx\,\delta x + dy\,\delta y}{ds} \right]_0^1 .$$

The brace under either integral may be set equal to zero to obtain the extremals. The first of these may be integrated and gives

$$gs + \text{const} = (\lambda g + gx - \tfrac{1}{2}\omega^2 y^2) \frac{dx}{ds} \, . \tag{7}$$

The conditions at the limits are

$$[(\lambda g + gx - \tfrac{1}{2}\omega^2 y^2)(dx\,\delta x + dy\,\delta y)]_0^1 = 0,$$

* This method of procedure which is almost always the most convenient in practice is admirably treated by de la Vallée Poussin in his *Cours d'analyse infinitésimale*, volume 2, pages 325–346.

and show that: *If the extremities of the chain are movable on fixed curves, the chain must be perpendicular to those curves.**

3. General discussion of the equilibrium. As equation (5) appears not to be integrable except into (7) which contains the arc and its derivative and is probably not further integrable, it will be well to begin with a general discussion of the equilibrium. In the first place we may state the result: *Two chains precisely alike except for different densities will have the same positions of equilibrium.* This may be seen at once from the equations (1) or (2). The physical reasoning which may be employed to reach the same conclusion is equally simple. It is merely necessary to cite the principle that if all the applied forces in a statical system be reduced in the same ratio, the system remains in equilibrium and the forces developed, that is, frictional forces, reactions at fixed constraints, and so on, are reduced in the same ratio. For similar reasons we may state: *The position of equilibrium will be the same in all fields which have the same value for the ratio $\omega^2 : g$.* In fact until the unit of time, which is of no consequence in statics, has been specified, it is only the ratio of ω^2 and g which can be said to be known. Therefore without loss of generality the equations may be simplified by setting $\rho = 1$ and $\omega^2/g = \kappa$.

The case most easily realized in practice is where the upper end of the string is fixed at a point in the axis of rotation and the other end allowed to hang free: the origin may then be taken at the point of support. It is well known in the theory of the conical pendulum that the angular velocity must attain a certain value relative to the length of the pendulum before the position of equilibrium will depart from the vertical. In case the pendulum is a uniform rod of length l, the total moment of the centrifugal forces about the point of support is $\frac{1}{3}\rho\omega^2 l^3 \sin\theta\cos\theta$, where θ is the angle of inclination. The moment of the gravitational forces is $\frac{1}{2}\rho g l^2 \sin\theta$. Hence for equilibrium

$$\sec\theta = \tfrac{2}{3}\omega^2 l/g; \qquad \text{and} \qquad \omega^2 l/g = \kappa l > \tfrac{3}{2}$$

is the necessary and sufficient condition that the position of stable equilibrium is not vertical.

It is readily shown that: *The chain will depart from the vertical if $\kappa l > \frac{3}{2}$.*

---

* The possible exception of a free end on (4′), which would satisfy the condition identically, proves later not to be an exception.

For if $\kappa l > \frac{3}{2}$, the facts concerning the rod show that there are positions in the neighborhood of the vertical, namely, all straight lines slightly inclined, for which $\int V\,ds$ is less than its value along the vertical subject to the condition of constant length. Hence there are variations of the integral which are negative, weak variations at that, and the vertical cannot be a position of stable equilibrium. Similarly in case the chain is attached to two points of the vertical axis separated by a distance $a < l$, it may be shown that the position of equilibrium will depart from the vertical if $\kappa(l + a) > 3$. Whether on the other hand the chain will hang in the vertical if these inequalities are not satisfied is a question to be taken up later.

If the chain has a free end the tension at that point is zero and conversely if the tension at any point is zero the chain may be considered as cut and having a free end at that point. As the introduction of new constraints never disturbs an existing equilibrium, the chain may be considered as pinned in the field indefinitely near the free end, and hence it is physically obvious that: *At the free end a chain must have the direction of the field, that is, must be perpendicular to the equipotential* (4′). Again if the tension be not zero, the second of the intrinsic equations (2) shows that: *A chain has a point of inflexion where and only where the chain has the direction of the field (except perhaps at the free end, if one exists)*. The curvature at the free end requires special investigation because here (2) gives $R = 0/0$.

To treat this question, transform (5) to a new origin of x so that the y-axis passes through the free end upon the parabola (4′) at a point $y = y_0$, and perform the integration which gives (7). Let s be measured positively from the free end up. The constant in (7) vanishes because s and the first member of (4′) vanish, and $\lambda = \frac{1}{2}\kappa y_0^2$. We have then

$$(7') \qquad s\,\frac{ds}{dx} = x + \frac{1}{2}\,\kappa(y_0^2 - y^2).$$

If we plot an element of the tangent $y = y_0(1 - \kappa\Delta x)$, we obtain at the end of this lineal element

$$\frac{ds}{dx} = \frac{\Delta x - \frac{1}{2}\kappa[\,y_0^2 - y_0^2(1 - \kappa\Delta x)^2]}{\sqrt{1 + \kappa^2 y_0^2}\quad \Delta x}.$$

Hence

$$\Delta\,\frac{ds}{dx} = \sqrt{1 + \kappa^2 y_0^2} - \frac{ds}{dx} = \frac{-\frac{1}{2}\,\kappa^3 y_0^2 \Delta x}{\sqrt{1 + \kappa^2 y_0^2}}.$$

Hence if we assume that as we plot the curve by the Cauchy-Lipschitz method of approximation, the limit approached by the polygon is the integral of the differential equation, and that the curvature of the integral may be obtained from the polygon,* we find

$$\frac{d^2s}{dx^2} = -\frac{1}{2}\frac{\kappa^3 y_0^2}{\sqrt{1+\kappa^2 y_0^2}}\,.$$

Inasmuch as

$$\frac{d^2s}{dx^2} = \frac{d\sec\tau}{dx} = \tan\tau\,\sec\tau\,\frac{d\tau}{ds}\cdot\frac{ds}{dx} = -\frac{\kappa y_0(1+\kappa^2 y_0^2)}{R}\,,$$

where R is the radius of curvature, it follows that

$$R = \frac{2(1+\kappa^2 y_0^2)^{\frac{3}{2}}}{\kappa^2 y_0^2}\,.$$

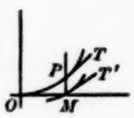
And if the radius of curvature of the lines of force

$$y = y_0 e^{-\kappa x}$$

defined by the differential equation $dy/dx = -\omega^2 y/g = -\kappa y$ be computed, it is found to be one-half of the value of R above. Hence: *The radius of curvature of the chain at its free end is neither zero nor infinite but twice the radius of curvature of the field at that point.*

---

* It is well to note precisely what this assumption amounts to. In the figure the curve OP represents the integral of the differential equation. The first step in the Cauchy-Lipschitz method is to construct the tangent OM. At the point M the tangent MT' is drawn. This is not tangent to the integral curve OP but to the neighboring integral curve through M. Thus the first two lines in the approximation are OM and MT'. The curvature of the integral OP is the limit of the inclination of PT divided by the arc OP as $P \doteq O$. In the text this limit has been replaced by the limit of the inclination of MT' divided by OM as $M \doteq O$. If now it be assumed, as it is, that at the point O, the curvature of the integral curve OP is not infinite, the difference between OP and OM is an infinitesimal of the third order at least, and PM is at least of the second order. The difference between the inclinations of PT and MT' will then also be of the second order at least Consequently in the limit OP may be replaced by OM and the inclination of PT may be replaced by that of MT'. Hence under the assumption that the curvature of OP at O is finite, the statement in the text is justified and the curvature may be found by the process indicated.



Suppose that the chain, whether it has a free end or not, lies in a curve which has maxima and minima of y. For brevity maxima and minima may be designated as extreme points. At each extreme point the tension is parallel to the x-axis, and as the tension is null at the free end (if there is one) we may include this point among the extreme points. As the chain is in equilibrium, it may be treated for the moment as a rigid body. The forces acting between two extreme points are the tensions at those points and the force of gravity parallel to the axis, and the elemental centrifugal forces perpendicular to the axis. These must form a system in equilibrium. Hence the centrifugal forces balance. But they have the form $\omega^2 y ds$ and hence between extreme points,

$$\int \omega^2 y \, ds = 0.$$

This is precisely the condition that the center of gravity of the chain shall be on the x-axis. We may therefore state: *The center of gravity of the arc of the chain between any two extreme points lies on the axis.* From this result some corollaries: 1° *The chain must cross the axis between successive extreme points.* 2° *The free end of the chain cannot lie on the axis.*

4. The form of the curves of equilibrium. To obtain a more precise idea of the curves of equilibrium, it is necessary to resort to a detailed discussion of the differential equation which defines them. It has been seen that the inflexions of the curves are tangent to the lines of force. Let us show that: *No inflexion can lie upon the axis.* Take the origin at the supposed inflexion. Equation (5) shows that if $y'' = 0$ and $y = 0$, then $y' = 0$. Now the axis of rotation is known to be a possible solution of the equation (5) of the second order with the initial conditions $y' = y = x = 0$. Hence, relying on the uniqueness of the solution, we see that no other solution with these initial conditions is possible, and therefore no inflexion can lie on the axis except in the case where the whole chain lies along the axis. The same conclusion may be derived directly by computing, by successive differentiations of (5), the values of all the higher derivatives $y^{(n)}$ at the point (0, 0) and noting that they all vanish.

Next consider the particular case of the chain with a free end and take (7′) as the defining equation. If this be written as

$$\frac{ds}{dx} = \frac{x + \frac{1}{2}(y_0^2 - y^2)}{s},$$

it appears that ds/dx cannot be infinite except at $s = 0$. When $s = 0$, however, it has been seen that ds/dx is the finite quantity $\sqrt{1 + \kappa^2 y^2}$. Hence: *In case the chain has a free end, the slope of the curve is everywhere finite.* Consider

$$\frac{ds}{dx} - 1 = \frac{x - s + \frac{1}{2}\kappa(y_0^2 - y^2)}{s}.$$

As $1 \leqq \sec \tau$, the numerator is always positive or zero. At the maximum and minimum points it vanishes. These points are therefore given by the equation

$$x - s + \tfrac{1}{2}\kappa(y_0^2 - y^2) = 0.$$

Here $x - s$ is zero at the free end and becomes increasingly negative. So at successive extreme points $y_0^2 - y^2$ must become increasingly positive, that is: *Each successive extreme point from the free end up is nearer the axis than the preceding.* Similarly it may be seen that: *At each successive intersection of the curve with the axis the inclination of the curve to the axis is less and is approaching zero as its limit.*

To return to the question of inflexions, we may note that it has not yet been shown that the chain actually changes its concavity at the inflexions. We shall therefore prove that: *The points of inflexion are of the first order, that is, $f'''(x) \neq 0$.* On differentiating (5) we find

$$y''' = -\frac{T(\kappa y' + y'')(1 + y'^2) + T(\kappa y + y')2y'y'' - (\kappa y + y')\,dT/dx}{T^2}.$$

As $dT/ds = -S$ and is finite, dT/dx is finite and y''' reduces to $-\kappa y'(1 + y'^2)/T$ at an inflexion and cannot vanish. We may next show that: *The curve must change its concavity between an extreme point and the next intersection with the axis.* For consider a minimum. As the curve is traced from this point up, the slope is positive and remains positive to the following maximum. The curve must therefore change its concavity while its slope is positive. As the slope at an inflexion is $-\kappa y$, the ordinate of the inflexion is negative. A maximum may be treated in the same manner. It may next be shown that: *There is only one inflexion between successive extreme points.* The easiest way to see this is to start at an inflexion, say on the positive side of the axis, and

trace the curve backward toward the free end. At the inflexion the slope is negative and becomes numerically less as the curve rises toward the maximum. But as the ordinates increase the slope of the field becomes negatively greater. Hence the maximum is reached before another inflexion. The slope of the curve then becomes positive and consequently cannot coincide with the slope of the field until the axis is crossed. When an inflexion on the negative side of the axis is reached the slope decreases as the curve rises whereas the slope of the field increases. Hence the minimum is reached before another inflexion, and so on.

It is readily shown that: *The maximum slope in each arch of the curve approaches zero as its limit* as well as the slope at the intersection with the axis. Hence: *The inflexions must approach the axis as a limit* although none of them lie upon it. There is no difficulty in seeing that: *The inclination of the curve is nowhere so great as at the free end and consequently the tension in the chain must be constantly increasing.* If we assume, as will be justified in the next section, that the ordinates of the curve approach zero as their limit, it appears that: *At least qualitatively the shape of the curve is represented by the graph of the Bessel's function J_0.* This will be further exemplified in article 5.

If the chain has not a free end, suppose that the y-axis be taken through the lowest intersection of the curve with the axis of x. Then equation (7) takes the form,

$$(s + \lambda \cos \tau_0) \frac{ds}{dx} = \lambda + x - \tfrac{1}{2}\kappa y^2,$$

where τ_0 is the inclination at the origin. According to the convention established above, s is to be measured positively if the curve starts off into either of the quadrants above the y-axis; in this case the discussion runs about as for the case of a free end. But if the curve starts out in either of the other two quadrants s has negative values. The form of the differential equation (5) is such that the curve cannot have any singular points until it meets the parabola (4′). In this case it cannot meet the parabola as there is no free end. We may therefore trace the curve back through negative values of s to a point where $s + \lambda \cos \tau_0$ vanishes and the curve becomes perpendicular to the axis. Obviously: *There is only one point at which the curve may have a tangent per-*

pendicular to the axis and this is situated below the lowest intersection with the axis, except in the limiting case where the tangent at the lowest point is itself perpendicular to the axis. If the y-axis be moved into coincidence with this tangent, the equation of the curve becomes

$$s\frac{ds}{dx} = x + \tfrac{1}{2}\kappa(y_0^2 - y^2),$$

which is identical in form with the equation (7′), in the case of a free end, except that here y_0 being the point of the parabola above the origin cannot be the ordinate of the point with the vertical tangent but must be greater. It is obvious by reasoning similar to that employed in the former case that: *The curve cannot rise so high as $y = y_0$. The ordinate will be doubly valued for all positive values of x and except for the vicinity of the vertical tangent each of the branches will behave qualitatively like the former curve.* Further details may be left to the reader.

5. Approximations in the differential equation. Let us next plot the curves by following successive tangent lines. If the curve has a free end, we may start at this end with equation (7′). This method avoids having to deal with the question of the value of T_0 or λ. In fact the length of the chain will be a consequence of the position of equilibrium from this point of view, instead of the reverse. To simplify matters and to have a definite parabola let us assume $\kappa = 2$. It may be noted that: *The constant κ is the reciprocal of the semi-latus rectum. The inequality $\kappa l > \frac{3}{2}$, if we may assume the necessity as well as sufficiency of the condition, shows that the least length of the chain is 3/2 the semi-latus rectum, that is, 3/2 the radius of curvature at the vertex of the parabola.* From the fact that the qualitative discussion has shown that the curve does not change its concavity between the free end and its first root it follows that: *No matter how near the vertex of the parabola the the free end may be, the length of the chain must be at least equal to the radius of curvature.* As there seems to be no other way of showing that this length must be half as great again as this radius of curvature, except by investigating the roots of the solution of the differential equation (a rather difficult and inaccessible problem), it may be interesting to plot one curve which starts very near the vertex.

Let us assume $y_0 = 0.05$, then at the start $y' = 0.1$. The following table of values may then be constructed.

Δs	s	x	y	$\log (ds/dx)$	$\tau = \sec^{-1} \dfrac{ds}{dx}$	$\Delta\tau$
0.	0.	0.	+0.05	0.0 021 607	5 42′ 38″	
0.05	0.05	0.04 975 165	0.04 502 485	0.0 019 449	5 25 6	17′ 32″
0.05	0.10	0.09 952 823	0.04 030 351	0.0 017 503	5 8 26	16 40
0.05	0.15	0.14 932 711	0.03 582 355	0.0 015 715	4 52 16	16 10
0.05	0.20	0.19 914 651	0.03 157 770	0.0 014 078	4 36 39	15 37
0.05	0.25	0.24 898 469	0.02 755 943	0.0 012 580	4 21 32	15 7
0.05	0.30	0.29 884 008	0.02 375 825	0.0 011 214	4 6 56	14 36
0.05	0.35	0.34 871 114	0.02 016 972	0.0 009 968	3 52 50	14 6
0.05	0.40	0.39 859 651	0.01 678 587	0.0 008 834	3 39 11	13 39
0.05	0.45	0.44 849 491	0.01 360 002	0.0 007 810	3 26 7	13 4
0.05	0.50	0.49 840 507	0.01 678 587	0.0 006 880	3 13 27	12 40
0.05	0.55	0.54 832 592	0.00 779 195	0.0 006 037	3 1 13	12 14
0.05	0.60	0.59 825 647	0.00 515 748	0.0 005 279	2 49 24	11 45
0.05	0.65	0.64 819 573	0.00 269 369	0.0 004 597	2 38 9	11 19
0.05	0.70	0.69 814 284	+0.00 039 431	0.0 003 985	2 27 15	10 54
0.05	0.75	0.74 809 698	−0.00 174 670			

The object of giving this table in extenso is to show the large number of figures necessary to get anything like a trustworthy result. As it is, the seconds are by no means determined; the differences of the log-secant are only about one unit in the last place given. The point at which the curve crosses the axis is found from the table to be $s = 0.718+$. Under our assumption that $\kappa = 2$, this length should exceed (probably by only an insignificant amount) the minimum length 0.75. As a seven-place table did not afford a ready accuracy sufficient to make the calculations when the increments of arc

are taken smaller and as such a calculation would be very tedious in any case, the calculation for increments of arc equal to 0.10 instead of 0.05 was carried through. This gave a length $s = 0.695-$, which is less than the other as was to be expected. Moreover the difference 0.023 or 0.024 between the two results is of such magnitude as to render it quite plausible that if the approximation were pushed to the limit the length would slightly exceed 0.75. It is indeed not difficult to obtain roughly an idea of the order of magnitude of the error introduced at each step of the approximation. If R be the radius of curvature and $\Delta\tau$ the angle between successive tangents, the departure of the tangent from the curve is $R(\Delta\tau)^2/2$. As $R = \Delta s/\Delta\tau$, this becomes $\Delta\tau\,\Delta s/2$. In the present problem $\Delta\tau$ is on the average about 1/250th of a radian. The individual errors in the fifteen approximations are cumulative, and hence the total error in the ordinate must be expected to be not far from 0.0015. As a matter of fact if the final ordinate had been greater by the amount 0.0018, the length of the chain from the free end to the intersection with the axis would have been in excess of 0.75. We may therefore fairly conclude that in case $\kappa = 2$ (and hence probably in all cases): *The condition $\kappa l > \frac{3}{2}$ is necessary and sufficient that the chain have a position of equilibrium other than the vertical.*

From the above discussion it appears that throughout the entire length of a curve which starts near the vertex of the parabola, the values of y' and y remain very small. As equation (5) may be written in the form

$$(5') \qquad y'' = -\frac{(\kappa y + y')(1 + y'^2)}{x + \frac{1}{2}\kappa(y_0^2 - y^2)},$$

it is further evident that at any rate when the curve reaches values of x large in comparison with y and y' — and this occurs very shortly — the difference between (5′) and

$$(5'') \qquad y'' = -\frac{(\kappa y + y')}{x}$$

becomes negligible. Of the two independent integrals of (5″) one is infinite at the origin and the other is $J_0(2\sqrt{\kappa x})$. Let us see how well

$$y = y_0 J_0(2\sqrt{\kappa x}) = y_0\left(1 - \kappa x + \frac{\kappa^2 x^2}{(1\cdot 2)^2} - \frac{\kappa^3 x^3}{(1\cdot 2\cdot 3)^2} + \cdots + (-1)^n \frac{\kappa^n x^n}{(n!)^2} + \cdots\right)$$

represents the curve of equilibrium. To begin with, the slope and radius of curvature at the origin are $-\kappa y_0$ and $2(1 + \kappa^2 y_0^2)^{3/2}/\kappa^2 y_0$. These are identical

with those of the curve of equilibrium. The form of (5″) shows that the inflexions of the Bessel's curve are tangent to the field. Finally, the first root of the Bessel's function is $2\sqrt{\kappa x} = 2.4+$ which in case $\kappa = 2$ becomes $0.72+$. This is remarkably close to 0.75 and represents a difference of ordinates of about 0.0013 between the curve of equilibrium and the approximation given by the Bessel's curve.

As the approximation is so extremely close in the only region where its accuracy would be under suspicion, we may safely conclude that it is safe to trust in regard to the rest of the curve. This would then establish the facts that the maximum ordinates of the curve of equilibrium approach zero as their limit. From a table of the roots of $J_0(x)$ we find

$$x = 2.40, \quad 5.52, \quad 8.65, \quad 11.79, \quad 14.93, \quad 18.07$$

for the successive roots. Hence in $J_0(2\sqrt{\kappa x})$ the successive roots are

$$\kappa x = 1.44, \quad 7.63, \quad 18.8, \quad 34.8, \quad 55.8, \quad 84.5.$$

The lengthening of the successive arcs is therefore established and a value which is certainly not far from the true one is obtained. If for the purposes of calculation, we assume that these are the true values of the intersections of a curve of equilibrium with the axis when the curve lies indefinitely near the axis, we see that the minimum lengths of a chain which shall cross the axis 0, 1, 2, 3, 4, 5 times are respectively

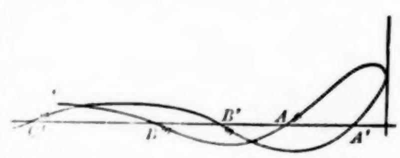
$$1.44/\kappa, \quad 7.63/\kappa, \quad 18.8/\kappa, \quad 34.8/\kappa, \quad 55.8/\kappa, \quad 84.5/\kappa.$$

As the interval between the roots of $J_0(x)$ rapidly approaches π, it is apparent that the additional amount of chain which must be provided in order to permit an additional intersection with the axis rapidly approaches $\pi\sqrt{l/\kappa} + \pi^2/4\kappa$ where l is the amount already provided.

6. Conclusions. Suppose that the parabola (4′) be drawn and the equilibrium curves which start from it at each point of the contour be sketched in. It is not difficult to see what are the possible positions of equilibrium for a curve of given length. Let the length be l. If l satisfies the inequality $1\frac{1}{2} < \kappa l < a$, where a is in the neighborhood of 8, the curve will hang entirely on one side of the vertical and will depart more and more from the vertical as κl lies nearer to a and further from $1\frac{1}{2}$. If l satisfies the inequality $a < \kappa l < \beta$, where β is in the neighborhood of 19, the curve may hang entirely on one side

of the vertical and in this case it will depart from the vertical more than the curves of the preceding case. Or the curve may cross the axis and will then recede from the vertical more and more as κl is nearer β and further from α. If l satisfies the inequality $\beta < \kappa l < \gamma$, where γ is in the neighborhood of 35 the curve may lie entirely on one side of the vertical and will depart from the vertical more than in the previous cases. Or it may cross the axis once and will depart further from the axis than in the preceeding case. Or it may cross the axis twice and depart from the vertical more and more as κl lies nearer γ and further from β. And so on. The discussion of the case of the chain with a free end has been pushed about as far as seems possible without actually integrating the differential equation (5).

The discussion of the case where both ends are fixed in the axis is by no means so simple. The reason is that in case there is a vertical tangent no approximate solution of (5) which shall adequately represent the curve both in the neighborhood of the vertical tangent and in the remoter regions is easily obtained. In the remoter regions, however, the approximation by means of $J_0(2\sqrt{\kappa x})$ holds also in this case. Hence the curve may be plotted in the neighborhood of the vertical tangent by the graphical method used in article 5 and then it may be pieced out with Bessel's curves. A rough sketch of the curve is given in the figure. The curve may be considered as fastened at any two of its intersections with the axis Undoubtedly there would be a definite set of inequalities to be satisfied in order that the curve should have 0, 1, 2, 3, $\cdots$ intersections with the axis (it being understood that a point of support on the axis does not count as an intersection). It is almost certain, however, that this set would not be of single entry as before, but of double entry. For the curve is double valued and each half obeys a law of decrease in ordinates and increase in length of arch similar to that found in the other case. Hence the minimum length of chain which may cut the axis once as between the points A and B' in the figure could hardly be expected to be the same as the minimum length of a chain which might cut the axis as between A' and B. In other words, if the total number of intersections were to be $n + m$, with n on one branch and m on the other, the minimum length would probably be $f(n, m)$ and not simply $f(n + m)$. But the investigation of this point would probably involve computations too laborious to be worth while.



In closing a word may be said of the behavior of the chain as the angular velocity is increased. The shape and properties of the curve are such that there is an absolute discontinuity between a possible equilibrium with different numbers of intersections with the axis. Hence if the angular velocity be gradually increased the curve cannot go gradually over from one type of equilibrium to another: the change is from one of less inclination with the axis to one of greater, and when the angular velocity has reached a certain value there will be an additional type of motion with one additional intersection with the axis but the chain will pass gradually over into it.*

In closing I should like to state that the problem of the rotating string was suggested to me, some long time since, by Professor Osgood, who had it from Professor Bôcher about a dozen years ago. The interesting feature to which my attention was called was the fact that the approximate solution by Bessel's functions was apparently valid for only a discrete set of values of κl; these are the values we have called $\kappa l = \alpha, \beta, \gamma, \cdots$. I hope that the foregoing discussion may have shown what the physical significance of this is and what the state of the chain is for intermediate values of κl.

Massachusetts Institute of Technology,
Boston, Mass., October, 1907.

---

* In an actual experiment with a medium brass chain about a meter long attached to an axis which could be rotated either by hand or by an electric motor, it was found difficult though not impossible to get curves which had none or one or more than five intersections with the vertical, easy to obtain those with two, three, four, or five intersections. In the first place note that extreme slowness of rotation is not particularly available owing to the large effects due to disturbances by air currents and other irregularities. Moreover if $l = 100$ and kl is about 100, which corresponds to five intersections, ω is only about five revolutions per second, which is a distinctly slow rotation. As this rotation is set up, the chain tends to climb up on itself and buckle, with the result that four or five intersections with the axis are likely to appear at the start and hence to persist. If the rotation is not much more rapid than five revolutions per second some of these intersections may easily be shaken out by sharp blows; if, however, the rotation is more rapid it becomes more and more difficult to disturb the type of motion (at any rate until such a rapidity is reached as will induce violent lateral vibrations in the chain and destroy the steadiness of the motion). This explains why the types with one or no intersections are hard to get with a rough apparatus. The difficulty in finding the types with six or more intersections appears to be due to the facts, that the chain is not a mathematical chain and that it must lie so near the axis that slight disturbances suffice to whip out one or more of the nodes. On the whole, the stability of each type of curve appears unexpectedly great.

THE CONTINUITY OF THE ROOTS OF AN ALGEBRAIC EQUATION

By J. L. Coolidge

The theorem that the roots of an algebraic equation whose term of highest degree has a non-vanishing coefficient are continuous functions of its coefficients is of fundamental importance, both in algebra and in geometry. The proofs usually given are rather long, and generally more closely allied in spirit to the theory of functions of a complex variable than to the elementary processes of algebra. For that reason a simple algebraic proof of this essential theorem seems to fill a real, if minute, gap in our algebraic theory. The object of this note is to give such a proof.

Theorem. *Given two algebraic equations of the same degree, so related that the coefficients of the first are constant and that of the highest power of the variable is not zero, while those of the second approach the corresponding coefficients of the first as limits; then the roots of the two equations, where each multiple root of order k is counted as k roots, may be put into such a one to one correspondence, that the absolute value of the difference of each two corresponding roots approaches zero.*

Let us write

$$f(x) \equiv a_0\, x^n + \cdots + a_n \equiv a_0(x - a_1) \cdots (x - a_n),$$

$$\phi(x) \equiv (a_0 + \Delta a_0)\, x^n + \cdots a_n + \Delta a_n \equiv (a_0 + \Delta a_0)(x - \beta_1) \cdots (x - \beta_n).$$

We assume explicitly that $a_0 \neq 0$ and that all of the quantities Δa_i approach zero. In particular we assume that $|\Delta a_0|$ is so small that for that, and all smaller values $|a_0 + \Delta a_0| > 0$.

Let us now reduce the roots of the equation $\phi(x) = 0$ by a_1. We shall get a new equation $\psi(x) = 0$, the coefficient of x^n being $(a_0 + \Delta a_0)$ while the constant term is $\phi(a_1)$. But since $f(a_1) = 0$, we shall have

$$\phi(a_1) = \Delta a_0\, a_1^n + \cdots \Delta a_n.$$

If $|\alpha_1| = 1$, $|\phi(\alpha_1)| \leqq (n+1)|\Delta a_k|$, where Δa_k is the largest Δa in absolute value.

$$\text{If} \qquad |\alpha_1| \neq 1, \qquad |\phi(\alpha_1)| \leqq \left| \Delta a_k \frac{|\alpha_1^{n+1}| - 1}{|\alpha_1| - 1} \right|,$$

and this approaches zero as a limit. The limit of $a_0 + \Delta a_0$ is a_0, hence the limit of the product of the roots of ψ namely $\pm \phi(\alpha_1)/(a_0 + \Delta a_0)$, is zero; hence one root, at least of $\psi(x) = 0$, approaches zero as a limit.

Let us, therefore, write

$$\beta_1 = \alpha_1 + \Delta\alpha_1,$$

where $\Delta\alpha_1$ approaches zero.

$$\begin{aligned} a_0x^n + \cdots a_n &\equiv (x - \alpha_1)(a_0x^{n-1} + b_1 x^{n-2} + \cdots b_{n-1}) \\ &\equiv (x - \alpha_1) f_1(x), \end{aligned}$$

$$\begin{aligned} &(a_0 + \Delta a_0)x_1^n + \cdots (a_n + \Delta a_n) \\ &\qquad \equiv \Big(x - (\alpha_1 + \Delta\alpha_1)\Big)\Big((a_0 + \Delta a_0)x^{n-1} + c_1x^{n-2} + \cdots c_{n-1}\Big) \\ &\qquad \equiv \Big(x - (\alpha_1 + \Delta\alpha_1)\Big)\phi_1(x). \end{aligned}$$

Multiplying out, and equating corresponding coefficients, we have

$$b_1 = a_0\alpha_1 + a_1, \qquad b_k = a_0a_1^k + a_1a_1^{k-1} + \cdots a_k,$$

$$c_1 = (a_0 + \Delta a_0)(a_1 + \Delta a_1) + (a_1 + \Delta a_1) = b_1 + \epsilon_1,$$

$$c_k = (a_0 + \Delta a_0)(a_1 + \Delta\alpha_1)^k + \cdots (a_k + \Delta a_k) = b_k + \epsilon_k,$$

where ϵ_k approaches zero with the Δ's.

Hence one root of $\phi_1(x) = 0$ lies infinitesimally near one root of $f_1(x) = 0$, or a second root of $\phi(x) = 0$ approaches a second root of $f(x) = 0$ as a limit.

By a repetition of this process our theorem is proved.

The foregoing reasoning is, of course, invalid in the case where $a_0 = 0$. It is, however, very easy to treat this case in a similar manner.

Let us assume that $a_n \neq 0$, for this may always be effected by reducing all the roots by a constant quantity—a process that does not in the least alter the validity of the conclusion.

We then write two new equations:

$$F(x) \equiv a_n x^n + \cdots \qquad + a_0 = 0,$$

$$\Phi(x) \equiv (a_n + \Delta a_n)x^n + \cdots \qquad +(a_0 + \Delta a_0) = 0.$$

To be perfectly general we shall assume that the first p coefficients in $f(x) = 0$, or the last p in $F(x) = 0$ are zero. Then $F(x) = 0$ has p roots equal to zero, and the remaining roots are the reciprocals of the roots of $f(x) = 0$. The non-vanishing roots of $\Phi(x) = 0$ are the reciprocals of those of $\phi(x) = 0$. Applying our theorem to $F(x) = 0$ and $\Phi(x) = 0$ we have:

THEOREM. *If two equations be so related that the coefficients of the first are constants, while each coefficient of the second approaches the corresponding coefficient of the first as a limit, and if the coefficients of the* p *highest powers of the unknown in the first be zero, then if each root of multiplicity* k *be counted as* k *roots, the second equation will have* p *roots whose absolute value will increase beyond all limit, and the remaining roots of the second may be put into such a one to one correspondence with the roots of the first, that each root of the second will approach the corresponding root of the first as a limit.*

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.,
JANUARY, 1908.

ON THE DIFFERENTIATION OF DEFINITE INTEGRALS

By Wm. F. Osgood

The object of this paper is to give a simpler proof of the theorem that

$$(A)\qquad \frac{d}{d\alpha}\int_a^b f(x,\alpha)dx = \int_a^b \frac{\partial f}{\partial \alpha}\,dx + f(b,\alpha)\,\frac{db}{d\alpha} - f(a,\alpha)\,\frac{da}{d\alpha}$$

than those that are current.

1. The Common Proofs. The theorem is usually proven by writing

$$\phi(\alpha) = \int_a^b f(x,\alpha)dx,$$

forming the difference:

$$\phi(\alpha+\Delta\alpha) - \phi(\alpha) = \int_{a+\Delta a}^{b+\Delta b} f(x,\alpha+\Delta\alpha)\,dx - \int_a^b f(x,\alpha)dx$$

$$= \left(\int_{a+\Delta a}^{a} + \int_a^b + \int_b^{b+\Delta b}\right) f(x,\alpha+\Delta\alpha)dx - \int_a^b f(x,\alpha)dx$$

$$= \int_a^b [f(x,\alpha+\Delta\alpha) - f(x,\alpha)]dx + \int_b^{b+\Delta b} f(x,\alpha+\Delta\alpha)dx$$

$$- \int_a^{a+\Delta a} f(x,\alpha+\Delta\alpha)dx,$$

and applying the theorems of mean value to these last three integrals; suitable assumptions being made about the continuity of the functions that enter. Cf., for example, Goursat-Hedrick, *Mathematical Analysis*, vol. I, §97.

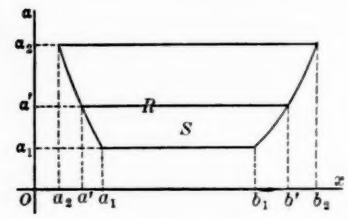
Another proof consists in changing the variable of integration so that the new limits of integration become constant:

$$t = \frac{x-a}{b-a}, \qquad x = (b-a)t + a,$$

$$\int_a^b f(x,\alpha)\,dx = (b-a)\int_0^1 f(x,\alpha)dt,$$

this latter case having been treated previously. Cf. Picard, *Traité d'analyse*, vol. I, chap. I, §18.

2. Critique of these Proofs. We will first state the theorem in detail in its simplest and most useful form. Let R be a region of the (x, α)-plane bounded by the right lines



$$\alpha = \alpha_1, \qquad \alpha = \alpha_2, \qquad (\alpha_1 < \alpha_2),$$

and the curves

$$a = \psi(\alpha), \qquad b = \omega(\alpha),$$

where each of the functions $\psi(\alpha)$ and $\omega(\alpha)$ shall be continuous, together with its first derivative, throughout the interval $\alpha_1 \leqq \alpha \leqq \alpha_2$, and where

$$\psi(\alpha) < \omega(\alpha).$$

The region R shall include its boundary.

The function $f(x, \alpha)$ shall be continuous at all points of R. Its partial derivative

$$\frac{\partial f}{\partial \alpha} = f_\alpha(x, \alpha)$$

shall exist and be continuous at all interior points of R, and this function, $f_\alpha(x, \alpha)$, shall remain finite throughout the interior of R.

Under the above conditions the integral

$$\int_a^b f(x, \alpha)\,dx$$

defines a function of α having a derivative given by the above formula, (A).

Turning now to the first of the proofs cited in §1 we see that if, in the neighborhood of a point $\alpha = \alpha'$ of the interval $\alpha_1 \leqq \alpha \leqq \alpha_2$, the functions $\psi(\alpha)$ and $\omega(\alpha)$ are both monotonic, the above transformation of the difference of the integrals, or a suitable modification of this transformation, is legitimate and the proof is sound. But there are several cases to consider when Δa and Δb change sign with $\Delta\alpha$.

If, in particular, the functions $f(x, \alpha), f_\alpha(x, \alpha)$ are continuous throughout a larger region R' lying between the same parallels $\alpha = \alpha_1$ and $\alpha = \alpha_2$ and containing all the interior and boundary points of R situated between these

lines in its interior, the single transformation of the difference given above will hold for both positive and negative values of Δa. But the theorem thus restricted is too narrow for all the ordinary applications of practice.

There are, then, even when $\psi(a)$ and $\omega(a)$ are both monotonic, several cases of the transformation of the difference of the above integrals to be considered.

If, however, one of the functions $\psi(a)$, $\omega(a)$ is not monotonic; for example, if

$$\omega(a) = (a - a')^3 \sin \frac{1}{a - a'}, \qquad a \neq a',$$
$$\omega(a') = 0,$$

there is trouble. The theorem is still true, but a special ϵ-proof is necessary.

The second proof is not open to this objection, but it turns out that it is necessary to assume the existence of the partial derivative of $f(x, a)$ with respect to x, and thus the proof does not apply to the theorem stated in the above generality.*

3. A New Proof. We can obtain a simple proof of the theorem as follows. Consider the double integral

$$(1) \qquad I = \int_S\int f_a(x, a)\,dS,$$

extended over the region

$$S: \qquad a \leqq x \leqq b, \qquad a_1 \leqq a \leqq a',$$

when $a_1 < a' \leqq a_2$. We can first evaluate I by means of the iterated integral

$$(2) \qquad I = \int_{a_1}^{a'} da \int_a^b f_a(x, a)\,dx.$$

Secondly, we can evaluate I by means of Green's Theorem:

$$\int_S\int f_a(x, a)\,dS = -\int_C f(x, a)\,dx.$$

* By means of this transformation, however, a simple proof can be given that

$$\int_a^b f(x, a)\,dx$$

is a continuous function of a if $f(x, a)$ is continuous in R, and a, b are continuous functions of a. The existence of a derivative with respect to x is here unnecessary.

Hence

$$I = -\int_{a_1}^{b_1} f(x, \alpha_1)\,dx - \int_{\alpha_1}^{\alpha'} f(b, \alpha)\,\frac{db}{d\alpha}\,d\alpha \tag{3}$$

$$+ \int_{a'}^{b'} f(x, \alpha')\,dx + \int_{\alpha_1}^{\alpha'} f(a, \alpha)\,\frac{da}{d\alpha}\,d\alpha.$$

Equating the two expressions (2) and (3) for I, dropping the accent against the α, and differentiating the equation thus resulting with respect to α, we obtain the theorem contained in the formula (A).

Instead of using the region S between the parallels $\alpha = \alpha_1$ and $\alpha = \alpha'$, we might have chosen an arbitrary value $\alpha_3 \neq \alpha'$. The two evaluations of I would have yielded the same final equation, the subscript 1 being merely replaced by 3. Thus, in particular, if $\alpha_3 = \alpha_2$, the excluded value $\alpha' = \alpha_1$ no longer presents an exception.

4. Generalizations. Corresponding to more general forms of Green's Theorem we obtain the theorem under consideration with less restrictive hypotheses. Thus if $f_\alpha(x, \alpha)$ is continuous within R, but does not remain finite on the boundary, the function $f(x, \alpha)$ still being assumed continuous in R, and if the surface integral (1) converges when extended over any region

$$S': \qquad a \leqq x \leqq b, \qquad \bar{\alpha}_1 \leqq \alpha \leqq \bar{\alpha}_2,$$

where $\alpha_1 < \bar{\alpha}_1 < \bar{\alpha}_2 < \alpha_2$, and if, moreover, the first integral to be evaluated in (2), namely:

$$\int_a^b f_\alpha(x, \alpha)\,dx,$$

converges uniformly in the interval $\bar{\alpha}_1 \leqq \alpha \leqq \bar{\alpha}_2$, the above proof will hold for all values of α' such that $\alpha_1 < \alpha' < \alpha_2$.

CAMBRIDGE, MASSACHUSETTS,
NOVEMBER 3, 1907.

NOTE ON THE CONVERGENCE OF A SEQUENCE OF FUNCTIONS OF A CERTAIN TYPE

By H. E. Buchanan and T. H. Hildebrandt

Theorems concerning uniform convergence of sequences and series of functions are numerous. Perhaps as well known as any is:

"If a sequence of continuous functions converges uniformly, the limit function is continuous."

The converse of this theorem does not, in general, hold. Harnack* proves a converse theorem for the case where the sequence of functions $f_n(x)$ is such that, for a definite n' and for every x, if $n \geqq n'$,

$$f_n(x) \leqq f_{n+1}(x).$$

Osgood† states a converse for the case

$$|f_n(x') - f_n(x)| \leqq M|x' - x|,$$

where M is a fixed positive quantity and x' and x are any two points of the interval considered. A converse of a slightly more general type than those of Harnack and Osgood is the following:

Theorem A. *If a sequence of monotonic nondecreasing‡ functions $f_n(x)$, $(n = 1, 2, 3, \cdots)$, converges to a function $f(x)$, continuous on the interval $ab (a \leqq x \leqq b)$, then $f(x)$ is a monotonic nondecreasing function, and the convergence is uniform over the interval ab.*

1. *The limit function is monotonic nondecreasing.* Suppose it were not. Then for certain two points $x_1, x_2 (a \leqq x_1 < x_2 \leqq b)$ we would have

$$f(x_1) > f(x_2).$$

Let

$$f(x_1) - f(x_2) = h.$$

* *Differential-und Integralrechnung*, p. 234.

† Annals of Mathematics, ser. 2, vol. 3, p. 139, 1902.

‡ Of course the theorem holds if everywhere "nondecreasing" is replaced by "nonincreasing."

Since $f_n(x)$ converges for all values of x on the interval $a \leqq x \leqq b$, we have

$$\underset{n}{L} f_n(x_1) = f(x_1)\,; \quad \underset{n}{L} f_n(x_2) = f(x_2)$$

that is, if there is given an $\epsilon > 0$, it is possible to find an $n_{\epsilon x_1}$ of such a nature that if $n \geqq n_{\epsilon x_1}$, we have

$$| f_n(x_1) - f(x_1) | \leqq \epsilon\,;$$

and an $n_{\epsilon x_2}$ of such a nature, that if $n \geqq n_{\epsilon x_2}$ we have

$$| f_n(x_2) - f(x_2) | \leqq \epsilon.$$

Suppose $\epsilon < \dfrac{h}{2}$ and choose n greater than or equal to $n_{\epsilon x_1}$ and $n_{\epsilon x_2}$. Then

$$| f_n(x_1) - f(x_1) | < \tfrac{1}{2}h \quad \text{and} \quad | f_n(x_2) - f(x_2) | < \tfrac{1}{2}h,$$

that is, $\qquad f_n(x_1) > f(x_1) - \tfrac{1}{2}h \quad \text{and} \quad f_n(x_2) < f(x_2) + \tfrac{1}{2}h,$

and therefore $\qquad f_n(x_1) - f_n(x_2) > 0.$

But by hypothesis $f_n(x)$ is monotonic nondecreasing, *i. e.*,

$$f_n(x_1) - f_n(x_2) \leqq 0.$$

We have then reached a contradiction, and therefore the hypothesis that $f(x)$ is not a monotonic nondecreasing function of x is invalid.

2. Since $f_n(x)$ and $f(x)$ are monotonic nondecreasing functions, their variations over any interval $x_1, x_2 (a \leqq x_1 < x_2 \leqq b)$ are the differences

$$f_n(x_2) - f_n(x_1) \quad \text{and} \quad f(x_2) - f(x_1).$$

Suppose $\quad V_n(x_1 x_2) = f_n(x_2) - f_n(x_1) \quad$ and $\quad V(x_1 x_2) = f(x_2) - f(x_1)$

Then $\qquad \underset{n}{L} V_n(x_1 x_2) = V(x_1 x_2).$

For, $\quad \underset{n}{L} V_n(x_1 x_2) = \underset{n}{L}(f_n(x_2) - f_n(x_1)) = f(x_2) - f(x_1) = V(x_1 x_2).$

3. From this we deduce the following

Lemma: If $\underset{n}{L} f_n(x) = f(x)$ under the conditions imposed in Theorem A, then it is possible to find for every $\epsilon > 0$ and for every ξ, $(a \leqq \xi \leqq b)$, a $\delta_{\epsilon\xi}$ and

an $n_{\epsilon\xi}$ such that, if

$$|x - \xi| \leqq \delta_{\epsilon\xi} \quad \text{and} \quad n \geqq n_{\epsilon\xi}$$

we have

$$|f_n(x) - f(x)| \leqq \epsilon.$$

For, $f(x)$ is continuous, that is, for every $\epsilon > 0$ we can find a $\delta_{\epsilon\xi}$, such that, if $|x - \xi| \leqq \delta_{\epsilon\xi}$ we have

$$|f(\xi) - f(x)| \leqq \tfrac{1}{5}\epsilon.$$

Hence
$$V(\xi - \delta_{\epsilon\xi}, \xi + \delta_{\epsilon\xi}) \leqq \tfrac{2}{5}\epsilon.$$

Now by 2,
$$\underset{n}{L}\, V_n(\xi - \delta_{\epsilon\xi}, \xi + \delta_{\epsilon\xi}) = V(\xi - \delta_{\epsilon\xi}, \xi + \delta_{\epsilon\xi})\,;$$

that is, for every $\epsilon > 0$ there exists an $n'_{\epsilon\xi}$ such that $n \geqq n'_{\epsilon\xi}$ implies:

$$|V_n(\xi - \delta_{\epsilon\xi}, \xi + \delta_{\epsilon\xi}) - V(\xi - \delta_{\delta\xi}, \xi + \delta_{\epsilon\xi})| \leqq \tfrac{1}{5}\epsilon,$$

and therefore
$$V_n(\xi - \delta_{\epsilon\xi}, \xi + \delta_{\epsilon\xi}) \leqq \tfrac{3}{5}\epsilon.$$

Moreover since

$$\underset{n}{L} f_n(\xi) = f(\xi)$$

for every $\epsilon > 0$ we can find an $n''_{\epsilon\xi}$ such that $n \geqq n''_{\epsilon\xi}$ implies:

$$|f_n(\xi) - f(\xi)| \leqq \tfrac{1}{5}\epsilon.$$

Now

$$|f_n(x) - f(x)| \leqq |f_n(x) - f_n(\xi)| + |f_n(\xi) - f(\xi)| + |f(\xi) - f(x)|.$$

If then we take

$$|x - \xi| \leqq \delta_{\epsilon\xi} \quad \text{and} \quad n \geqq n_{\epsilon\xi} \geqq \begin{cases} n'_{\epsilon\xi} \\ n''_{\epsilon\xi} \end{cases}$$

we have

$$|f_n(x) - f(x)| \leqq V_n(\xi - \delta_{\epsilon\xi}, \xi + \delta_{\epsilon\xi}) + \tfrac{1}{5}\epsilon + \tfrac{1}{5}\epsilon \leqq \epsilon.$$

4. We now proceed to the proof of Theorem A. Suppose there is given an $\epsilon > 0$. We can then find by our Lemma, for every point ξ on the interval $a \leqq \xi \leqq b$, a $\delta_{\epsilon\xi}$ and an $n_{\epsilon\xi}$ such that if x lies in the $\delta_{\epsilon\xi}$ vicinity of ξ and $n \geqq n_{\epsilon\xi}$ we have

$$|f_n(x) - f(x)| \leqq \epsilon.$$

The $\delta_{\epsilon\xi}$ vicinities will form an infinity of segments covering the interval $a \leqq x \leqq b$. Hence by the Heine-Borel Theorem,* it is possible to select a finite number of these vicinities which will cover the interval. To these vicinities there will correspond a finite number of $n_{\epsilon\xi}$ as determined by the Lemma. Of these $n_{\epsilon\xi}$ we choose n_ϵ, the largest. This n_ϵ is evidently independent of the point ξ. Then for every $n \geqq n_\epsilon$ and $a \leqq x \leqq b$ we shall have

$$|f_n(x) - f(x)| \leqq \epsilon,$$

that is, the convergence is uniform.

5. Denote by $V_n(x)$ and $V(x)$ the total variations (in the sense of Jordan) of $f_n(x)$ and $f(x)$ from a to x. Then we can state the following theorem :†

THEOREM B. *If* $\underset{n}{L} f_n(x) = f(x)$ *and* $\underset{n}{L} V_n(x) = V(x)$ *for every* x *on the interval* $a \leqq x \leqq b$, *and* $f(x)$ *is continuous, then the convergence of* $f_n(x)$ *to* $f(x)$ *is uniform for* $a \leqq x \leqq b$.

The proof of this theorem can be made on the same lines as the proof of Theorem A.

CHICAGO, ILL.,
DECEMBER, 1907.

---

* *Annales de l'Ecole Normale Superieure*, vol. 12, p. 51. Veblen and Lennes, *Infinitesimal Analysis*, p. 34.

† This theorem was suggested by Professor Osgood.

ON THE INVERSE PROBLEM OF THE CALCULUS OF VARIATIONS *

By G. A. Bliss

In the second volume of the Annals,† Prof. Osgood has shown in a very interesting way why an arc

$$y = y(x), \qquad x_0 \leqq x \leqq x_1,$$

which gives a minimum (or a maximum) value to an integral of the form

$$(1) \qquad \int g(x, y, y_x)\,dx,$$

must be an extremal, that is, a solution of the differential equation

$$(2) \qquad g_y - \frac{d}{dx} g_{y'} = 0.$$

This equation is of the second order, and its solutions are therefore a two-parameter family of curves.

In the so-called inverse problem of the calculus of variations the two-parameter family of solutions of equation (2) is given in advance, and it is required then to find the integrals of the form (1) for which these curves are the extremals. In connection with some problems in differential geometry Darboux‡ has found that there is an infinity of such integrals, and showed how they may be determined. For more general cases the corresponding problems are discussed in an interesting article by Hirsch.§

In the present paper it is proposed to apply the method of Darboux to integrals in the form

$$(3) \qquad \int f(x, y, \tau)\sqrt{x_t^2 + y_t^2}\,dt.$$

* The first part of this paper, §§1, 2, was presented before The Chicago Section of the American Mathematical Society, April 22, 1905, under the title "The inverse problem of the calculus of variations in parametric representation."

† Ser. 2, vol. 2, p. 107; 1901.

‡ *Leçons sur la théorie générale des surfaces*, vol. 3, §§604–605.

§ *Mathematische Annalen*, vol. 49 (1897), p. 49.

The arcs along which the integral is taken are supposed to be represented in the form

$$(4) \qquad x = x(t), \quad y = y(t), \qquad t_0 \leqq t \leqq t_1,$$

and τ is the angle between the tangent and the x-axis defined by the equations

$$\cos\tau = \frac{x_t}{\sqrt{x_t^2 + y_t^2}}, \quad \sin\tau = \frac{y_t}{\sqrt{x_t^2 + y_t^2}},$$

or, what is the same thing, by

$$(5) \qquad \tau = \tau_0 + \int_{t_0}^{t} \frac{x_t y_{tt} - x_{tt} y_t}{x_t^2 + y_t^2}\, dt,$$

where τ_0 is the initial value of τ at the point $t = t_0$. The integral (3) evidently reduces to (1) when

$$f(x, y, \tau) = g(x, y, \tan\tau)\cos\tau.$$

The geometrical advantage in using an integral of the form (3) instead of (1) is due to the fact that the arcs over which it is taken may go in any direction, while in the form in which Darboux considered the problem difficulties are always encountered when the tangent to the curve becomes parallel to the y-axis, that is, when y_x becomes infinite. For the inverse problem the integral (3) is also decidedly more convenient than the well-known integral

$$(6) \qquad \int_{t_0}^{t_1} F(x, y, x_t, y_t)\, dt$$

which is usually used when the curves are taken in parametric representation. This is because the function f is determined by integrating a single linear partial differential equation instead of the system of such equations which present themselves when the solution of the problem for the integral (6) is attempted.

The first two sections below are devoted to the derivation for the integral (3) of the so-called Euler equation corresponding to equation (2), and to the determination of the integrals in the form (3) which have a given family of extremals. The integrals which Darboux gives and those which are found in §2 involve two arbitrary functions. In §3 it is shown what further conditions,

besides the prescription of the extremals, suffice to determine the integral uniquely, and in §4 the results of the preceding sections are applied to the determination of the integrals which are minimized by straight lines.

1. Euler's equation. The derivation of Euler's equation for the extremals of the integral (3) has been made in a previous paper,* but for the sake of completeness it will be repeated here. An arc E in the form (4) is supposed to give to the integral a smaller value than that given by any other curve in its neighborhood with the same endpoints. A variation V of E in the form

$$(V) \qquad X = x(t) + a\xi(t), \qquad Y = y(t) + a\eta(t), \qquad t_0 \leqq t \leqq t_1,$$

where ξ and η are two functions of t vanishing at t_0 and t_1, can be made to lie as near as is desired to E by taking the parameter a sufficiently small, and for $a = 0$ the arc V coincides with E itself. It is evident then that the integral

$$I(a) = \int_{t_0}^{t_1} f(X, Y, T)\sqrt{X_t^2 + Y_t^2}\, dt$$

is a function of a which must have a minimum, and consequently the derivative $\frac{dI}{da}$ must vanish, for $a = 0$.

The derivative can be readily calculated when it is noticed that

$$X_a = \xi, \quad X_{ta} = \xi_t, \quad Y_a = \eta, \quad Y_{ta} = \eta_t,$$

and that

$$T_a = \frac{X_t\eta - Y_t\xi}{X_t^2 + Y_t^2}$$

since

$$\cos T = \frac{X_t}{\sqrt{X_t^2 + Y_t^2}}, \qquad \sin T = \frac{Y_t}{\sqrt{X_t^2 + Y_t^2}}.$$

Then

$$\left[\frac{dI}{da}\right]_{a=0} = \int_{t_0}^{t_1}\left\{f_x\xi + f_y\eta + (f\cos\tau - f_\tau\sin\tau)\xi_s + (f\sin\tau + f_\tau\cos\tau)\eta_s\right\}ds,$$

where

$$ds = \sqrt{x_t^2 + y_t^2}\, dt,$$

---

* A generalization of the notion of angle, *Transactions of the American Mathematical Society*, vol. 7 (1906), p. 185.

or by a partial integration

$$(7)\qquad \left[\frac{dI}{da}\right]_{a=0} = \int_{t_0}^{t_1}\left\{\left[f_x - \frac{d}{ds}(f\cos\tau - f_\tau \sin\tau)\right]\xi + \left[f_y - \frac{d}{ds}(f\sin\tau + f_\tau\cos\tau)\right]\eta\right\}ds.$$

This expression must be zero however ξ and η are chosen, and by the usual argument in the calculus of variations it follows that along the arc E the two expressions

$$P = f_x - \frac{d}{ds}(f\cos\tau - f_\tau\sin\tau), \qquad Q = f_y - \frac{d}{ds}(f\sin\tau + f_\tau\cos\tau),$$

must vanish. These conditions upon E are not independent, for the relation

$$(8)\qquad P\cos\tau + Q\sin\tau = 0$$

holds between P and Q, and one verifies readily that the vanishing of the two expressions P and Q is equivalent to the vanishing of the expression T defined by the equation

$$(9)\qquad P\sin\tau - Q\cos\tau = T.$$

In fact from (8) and (9)

$$P = T\sin\tau, \quad Q = -T\cos\tau.$$

From either of these equations the value of T is easily calculated.

Along any arc E which minimizes the integral (3) *the expression*

$$(10)\qquad T(x, y, \tau, \tau_s) = f_x\sin\tau - f_y\cos\tau + f_{x\tau}\cos\tau + f_{y\tau}\sin\tau + (f + f_{\tau\tau})\tau_s$$

must vanish identically. The equation

$$(11)\qquad T(x, y, \tau, \tau_s) = 0$$

is called the Euler equation for the integral (3).

2. Determination of the integrals which have given extremals. Consider a family of curves

$$(12)\qquad x = x(t, u, v), \quad y = y(t, u, v),$$

involving the two parameters u and v. From equation (5) the direction angle τ is also a function of t, u, v,

$$\tau = \tau(t, u, v), \tag{13}$$

and it will be supposed that the three equations (12) and (13) have single-valued solutions

$$t = t(x, y, \tau), \qquad u = u(x, y, \tau), \qquad v = v(x, y, \tau) \tag{14}$$

for all points (x, y) in a certain region R of the xy-plane, and for arbitrary values of τ. In other words through each point of R there passes one and but one of the curves (12) in a given direction τ. All six of the functions are supposed to be continuous and have continuous first derivatives.

A family of curves of this sort will consist of solutions of a system of differential equations of the form

$$x_s = \cos\tau, \quad y_s = \sin\tau, \quad \tau_s = \phi(x, y, \tau). \tag{15}$$

For from equation (13) and the expression for s_t, the curvature $\tau_s = \dfrac{\tau_t}{s_t}$ is seen to be a function of t, u, v, and by equations (14) of x, y, τ. Conversely, when the three equations (15) are given, it follows from the well-known theory of differential equations that there is but one two-parameter family of curves which satisfy them. The assumption that the curves (12) are the extremals of the integral (3), and the assumption that the extremals must satisfy a given system of differential equations in the form (15), are therefore equivalent.

If the extremals of the integral are to be the solutions of the system (15), then from (11) the function f must satisfy the equation

$$f_x \sin\tau - f_y \cos\tau + f_{x\tau}\cos\tau + f_{y\tau}\sin\tau + (f + f_{\tau\tau})\phi = 0 \tag{16}$$

identically in x, y, τ. This is a partial differential equation of the second order for f and not easily solvable by well-known methods. But if it is differentiated partially for τ, another differential equation is found,

$$M_x \cos\tau + M_y \sin\tau + M_\tau \phi + M\phi_\tau = 0, \tag{17}$$

which is linear in the function

$$M = f + f_{\tau\tau} \tag{18}$$

and its derivatives of the first order.

The equations (17) and (18) can now be solved by a method considerably simpler than that of Darboux since it avoids the use of the theory of partial differential equations. By means of equations (12) and (13), $M(x, y, \tau)$ may be thought of as a function of t, u, v. From equation (17), then,

$$\frac{dM}{dt} + M\phi_\tau \sqrt{x_t^2 + y_t^2} = 0,$$

since $\tau_s = \phi$ and $M_t = M_s\sqrt{x_t^2 + y_t^2}$. Therefore

$$M = G(u, v)e^{-\int \phi_\tau \sqrt{x_t^2 + y_t^2}\, dt}, \tag{19}$$

where x, y, τ in ϕ_τ are supposed to be replaced by their values in terms of t, u, v, and the integral in the exponent stands for some particular function of t, u, v whose derivative for t is $\phi_\tau\sqrt{x_t^2 + y_t^2}$. The function $G(u, v)$ is an arbitrary function of u and v. The determination of M as a function of t, u, v in this manner means that it is also determined as a function of x, y, τ on account of equations (14).

It follows of course from the foregoing discussion that the function $M(x, y, \tau)$ defined by equation (19) satisfies the differential equation (17), but this can also be readily verified by direct substitution. The first member of (17) becomes simply

$$\begin{aligned} &M_u(u_x \cos\tau + u_y \sin\tau + u_\tau\phi) \\ &+ M_v(v_x \cos\tau + v_y \sin\tau + v_\tau\phi) \\ &+ M_t(t_x \cos\tau + t_y \sin\tau + t_\tau\phi) + M\phi_\tau, \end{aligned}$$

which vanishes since u and v are constant along an extremal and

$$M_t(t_x \cos\tau + t_y \sin\tau + t_\tau\phi) = M_t\frac{dt}{ds} = -M\phi_\tau,$$

from equation (19).

It should perhaps be mentioned that the functions u and v above are not the only ones which might be used in the determination of M. If u and v are any two independent solutions of the abbreviated equation

$$\Phi_x \cos\tau + \Phi_y \sin\tau + \Phi_\tau\phi = 0, \tag{20}$$

and if ψ is a particular solution of (17), then it is true that $\frac{M}{\psi}$, u, v, thought

of as functions of the four variables x, y, τ, M, are independent solutions of the auxiliary linear differential equation

$$\Phi_x \cos\tau + \Phi_y \sin\tau + \Phi_\tau \phi - \Phi_M M\phi_\tau = 0.$$

From the theory of linear equations* it follows that the most general solution $M(x, y, \tau)$ of equation (17) is found by solving the equation

$$\frac{M}{\psi} = G(u, v) \tag{21}$$

for M. The functions u and v from equations (12), and the exponential factor in equation (19) are particular solutions u, v, ψ of equations (20) and (17), respectively.

It remains to derive the value of f from that of M by means of the equation (18). This is a differential equation of the second order with respect to τ, having as its most general solution†

$$f = \int_0^\tau \sin(\tau - \lambda) M(x, y, \lambda) d\lambda + A(x, y)\cos\tau + B(x, y)\sin\tau, \tag{22}$$

where A and B are arbitrary functions of x and y. It is not certain that the value of f just given satisfies the Euler equation (16). In fact if the expression (22) is substituted in (16), the equation

$$\frac{\partial A}{\partial y} - \frac{\partial B}{\partial x} = \int_0^\tau (M_x \cos\lambda + M_y \sin\lambda) d\lambda + \phi(x, y, \tau) M(x, y, \tau) \tag{23}$$

is found. The function on the right is independent of τ since its derivative for τ is the expression (17), and consequently equation (23) may be written

$$\frac{\partial A}{\partial y} - \frac{\partial B}{\partial x} = \phi(x, y, 0) M(x, y, 0). \tag{24}$$

If A_0 and B_0 are any two particular solutions of this last equation, then it can readily be shown that the differences $A - A_0$, $B - B_0$, in which (A, B) is any other solution, satisfy the equation

$$\frac{\partial(A - A_0)}{\partial y} = \frac{\partial(B - B_0)}{\partial x},$$

* See for example Jordan, *Cours d'Analyse*, vol. 3 (1896), p. 314.

† For a discussion of linear differential equations with constant coefficients, see Murray, *Differential Equations*, chap. VI.

and consequently that the most general solution of (24) is

$$A = A_0 + \frac{\partial \theta}{\partial x}, \qquad B = B_0 + \frac{\partial \theta}{\partial y},$$

where θ is an arbitrary function of x and y.

Hence *the most general function* $f(x, y, \tau)$ *for which the integral* (3) *has a given set of extremals* (12), *is given by the equation*

$$(25) \quad f = \int_0^\tau \sin(\tau - \lambda)\, M(x, y, \lambda)\, d\lambda + \left(A_0 + \frac{\partial \theta}{\partial x}\right) \cos \tau + \left(B_0 + \frac{\partial \theta}{\partial y}\right) \sin \tau.$$

In this equation

$$M(x, y, \tau) = G(u, v) e^{-\int \phi \tau^{\prime} \sqrt{x_t^2 + y_t^2}\, dt},$$

where $G(u, v)$ *is an arbitrary function of* u *and* v, *and the arguments* t, u, v *in the expression on the right can be replaced by their values* (14) *in* x, y, τ. *The pair of functions,* $A_0(x, y)$ *and* $B_0(x, y)$, *are a particular solution of the differential equation*

$$\frac{\partial A}{\partial y} - \frac{\partial B}{\partial x} = \phi(x, y, 0)\, M(x, y, 0),$$

and θ *is an arbitrary function of* x *and* y.

3. Further conditions determining f uniquely. It is natural to ask at this point what further conditions on the function f would suffice to determine it uniquely. In order to determine the function $G(u, v)$ suppose that the region R has a simply closed curve

$$(B) \qquad x = g(\alpha), \quad y = h(\alpha)$$

as its boundary, which nowhere satisfies the Euler differential equation (11). It is evident that no extremal can be a closed curve in R, otherwise after a complete circuit the same values of x, y, τ would correspond to different values of t, which is contrary to the hypotheses upon the functions (12). It will be supposed, furthermore, that for any fixed values of u, v corresponding to a point of R and a direction through that point, the functions (12) of t can be continued until the extremal encounters the boundary curve B. No extremal E can cross the curve B at a point of tangency with B. For two curves tangent to each other cannot cross unless they have the same curvature,

and by hypothesis B can never satisfy the equation (11) which defines the curvature τ_s of E. It follows then that every extremal of the set (12) crosses the boundary curve B without being tangent to it, at one point at least.

Let now the values of $M(x, y, \tau)$ be assigned for any point (x, y) on the curve B and for all values of τ. The only restriction upon the values assigned to M, aside from the restriction that $M(a, \tau)$ as a function of a and τ shall be continuous and have continuous first derivatives, will be that the values of the quotient $\frac{M}{\psi}$ shall be the same at all points which the extremal E has in common with the curve B. Thus if the region R were a circle and the extremal E a straight line, $\frac{M}{\psi}$ would be the same for the values of x, y, τ on the straight line at both its intersection points with the circle. This restriction is necessary because from equation (21), $\frac{M}{\psi}$ is to be everywhere a function of u and v alone.

It follows now easily that for every value of u and v defining an extremal E in the region R, the value of $G(u, v)$ is completely determined. For the extremal E intersects B at least at one point, and there the value of M has been assigned, so that the value of G is completely determined by equation (21).

The analytical expression for $G(u, v)$ can be determined in an actual case by solving the equations

$$(26) \qquad u = u[g(a), h(a), \tau], \quad v = v[g(a), h(a), \tau]$$

for a and τ and substituting the results so found in the expression

$$\frac{M(a, \tau)}{\psi[g(a), h(a), \tau]}.$$

The values of u, v defined by equations (26) at a point b where an extremal E crosses B, are those belonging to E. The functional determinant

$$\begin{vmatrix} u_x g_a + u_y h_a & u_\tau \\ v_x g_a + v_y h_a & v_\tau \end{vmatrix} = \begin{vmatrix} u_x & u_\tau \\ v_x & v_\tau \end{vmatrix} g_a + \begin{vmatrix} u_y & u_\tau \\ v_y & v_\tau \end{vmatrix} h_a,$$

is different from zero at the intersection point. For along the extremal E

$$u_x \cos\tau + u_y \sin\tau + u_\tau \phi = 0,$$
$$v_x \cos\tau + v_y \sin\tau + v_\tau \phi = 0,$$

and therefore

$$\cos\tau : \sin\tau : \phi = \begin{vmatrix} u_y & u_\tau \\ v_y & v_\tau \end{vmatrix} : \begin{vmatrix} u_\tau & u_x \\ v_\tau & v_x \end{vmatrix} : \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix},$$

so that the functional determinant above is

$$(-\cos\tau g_a + \sin\tau h_a)\left\{ \begin{vmatrix} u_x & u_\tau \\ v_x & v_\tau \end{vmatrix}^2 + \begin{vmatrix} u_y & u_\tau \\ v_y & v_\tau \end{vmatrix}^2 \right\}.$$

This cannot vanish at a point b where E crosses B, since at such a point B and E are never tangent. Consequently from the theory of implicit functions* the solutions of the equations (26) are continuous functions of u and v with continuous first derivatives in the neighborhood of the values u, v which define any extremal E in the region R.

The only arbitrary part of f besides $G(u, v)$ is the function $\theta(x, y)$, and the values of θ can be determined by prescribing the transversals to any one-parameter family of curves C which simply cover the region R, that is, which cover the region in such a way that through each point there passes one and but one curve. In such a field of curves the direction angle τ is a function of x and y, and the transversal curves in the field are by definition the curves of a one-parameter family whose direction angles τ' satisfy the equation

$$(27) \qquad (f\cos\tau - f_\tau \sin\tau)\cos\tau' + (f\sin\tau + f_\tau\cos\tau)\sin\tau' = 0.\dagger$$

If the transversals are prescribed as a one-parameter family of curves, $w(x, y) = \text{constant}$, simply covering the field and nowhere tangent to the original family, then τ' is given as a function of x and y by the equations

$$\cos\tau' = \frac{w_x}{\sqrt{w_x^2 + w_y^2}}, \qquad \sin\tau' = \frac{w_y}{\sqrt{w_x^2 + w_y^2}},$$

and is different at every point from $\tau(x, y)$. By substituting the value (25) in equation (27) it follows that the function θ must satisfy the differential equation

$$(28) \qquad \frac{\partial\theta}{\partial x}\cos\tau' + \frac{\partial\theta}{\partial y}\sin\tau' + Q(x, y) = 0,$$

---

* Goursat-Hedrick, *A Course in Analysis*, vol. 1, §25.

† Bliss, A new form of the simplest problem of the calculus of variations, *Transactions Amer. Math. Soc.*, vol. 8 (1907), p. 413.

where

$$Q(x, y) = A_0 \cos\tau' + B_0 \sin\tau' + \int_0^{\tau} \sin(\tau' - \lambda) M(x, y, \lambda) d\lambda.$$

The most general solution of equation (28) can readily be determined if a particular solution $\chi(x, y)$ is known. For the function $w(x, y)$ is a solution of the abbreviated equation

$$\frac{\partial w}{\partial x} \cos\tau' + \frac{\partial w}{\partial y} \sin\tau' = 0, \tag{29}$$

and w and $\theta - \chi$ are two independent integrals of the auxiliary equation

$$\Phi_x \cos\tau' + \Phi_y \sin\tau' - Q\Phi_\theta = 0.$$

The most general solution of (28) is therefore determined by the equation

$$\theta - \chi = H(w),$$

where H is an arbitrary function of w.

Along any particular curve C_0 belonging to the original field, w varies monotonically, for on account of equation (29) it follows that

$$\frac{dw}{ds} = w_x \cos\tau + w_y \sin\tau \neq 0.$$

If this inequality were not true then at some point the ratios $\cos\tau : \sin\tau$ and $\cos\tau' : \sin\tau'$ would be equal and the transversal at that point would be tangent to C_0. Consequently s is also a monotonic function of w. If now the values of f as a function of s are given along C_0, then the arbitrary function H is determined except for an additive constant as a function of s by the equation

$$f(s) = \int_0^{\tau} \sin(\tau - \lambda) M(x, y, \lambda) d\lambda + (A_0 + \chi_x)\cos\tau + (B_0 + \chi_y)\sin\tau + \frac{dH}{ds},$$

and therefore, from what has just been said, as a function of w.

The results of this section are as follows:

If the boundary of the region R consists of a simply closed curve B which at no point satisfies the Euler equation, then the arbitrary function $G(u, v)$ is completely determined when the values which $M = f + f_{\tau\tau}$ takes at points of the boundary B for all values of τ, are assigned. Furthermore, the function $\theta(x, y)$ is determined, except for an additive constant, when, for any one-

parameter family of curves C which simply cover the region R, the family of transversal curves is assigned and the values of f along a particular curve C_0 are given. The function f is then uniquely determined in all that portion of the region R which is swept out by the transversal curves which intersect C_0.

4. The integrals having straight lines as extremals. As an application of the results of the preceding section consider the integrals which have as extremals the straight lines of the plane

$$x = t\cos v + u\sin v, \qquad y = t\sin v - u\cos v.$$

Here u is the distance of the line from the origin, and v is the angle which it makes with the positive x-axis. The solutions of these equations for t, u, v in terms of x, y, τ are

$$t = x\cos\tau + y\sin\tau, \quad u = x\sin\tau - y\cos\tau, \quad v = \tau,$$

and the differential equations (15) are simply

$$\frac{dx}{ds} = \cos\tau, \quad \frac{dy}{ds} = \sin\tau, \quad \frac{d\tau}{ds} = 0.$$

The value of $M(x, y, \tau)$ from equation (19) is

$$M = G(x\sin\tau - y\cos\tau, \tau),$$

the value $\psi = 1$ being an admissable value of ψ. Finally A_0 and B_0 can be taken equal to zero since the function $\phi(x, y, 0)$ in (24) vanishes. *The most general function f whose integral (3) has straight lines as extremals is therefore*

$$(30) \quad f = \int_0^\tau \sin(\tau - \lambda)G(x\sin\lambda - y\cos\lambda, \lambda)d\lambda + \theta_x\cos\tau + \theta_y\sin\tau,$$

where G is an arbitrary function of λ and $u = x\sin\lambda - y\cos\lambda$, and θ is an arbitrary function of x and y.

If the values of $M(x, y, \tau)$ at points of a circle

$$x^2 + y^2 = a^2$$

are assumed to be always equal to unity, then from the results of the last section $G(u, v)$ must be identically equal to unity for all values of u and v corresponding to straight lines which cut the circle. If the straight lines parallel to the x-axis are assumed to have the straight lines parallel to the

y-axis as transversals, then in equation (28) $\tau = 0$, $\tau' = \pi/2$, and the equation reduces to

$$\frac{\partial \theta}{\partial y} = 0.$$

Consequently θ is a function of x alone. If the values of f on the line $y = 0$ are assumed to be all equal to unity, then from (30)

$$1 = \theta_x, \quad \theta = x + c,$$

and it follows that the function f is everywhere equal to unity.

The integral

$$\int \sqrt{x_t^2 + y_t^2}\, dt$$

is therefore the most general integral (3) *for which*:

a) *straight lines are extremals,*

b) *the function $M(x, y, \tau)$ equals unity for values (x, y) on a circle with its center at the origin and for all values of τ,*

c) *the family of straight lines $y = c$ has as its family of transversals the lines $x = c$,*

d) *on the line $y = 0$ the values of $f(x, 0, 0)$ are unity.*

One of the problems which Darboux solved * was the determination of the form which the length integral on a surface has when the geodesics on the surface are represented in the plane of the parameters of the surface as straight lines. If the variables x and y are the parameters, the length integral will be an integral (3) in which

$$f(x, y, \tau) = \sqrt{E \cos^2\tau + 2F \cos\tau \sin\tau + G \sin^2\tau}, \tag{31}$$

E, F, and G being functions of x and y only. The problem is then to determine the functions f which are simultaneously of the forms (30) and (31). For (31) the value of M is

$$M = \frac{EG - F^2}{f^3},$$

and this must be a function of τ and $u = x \sin\tau - y \cos\tau$. The same must be true of $M^{-2/3}$, that is,

$$\frac{E \cos^2\tau + 2F \cos\tau \sin\tau + G \sin^2\tau}{(EG - F^2)^{2/3}} = \Phi(u, \tau). \tag{32}$$

* Loc. cit., §607.

The functions E, F, and G can be determined from this identity if $\Phi(u, \tau)$ is any homogeneous expression of the second degree in u, $\cos\tau$, $\sin\tau$. For since $u = x\sin\tau - y\cos\tau$, Φ may be arranged in powers of $\cos\tau$ and $\sin\tau$,

$$\Phi(u, \tau) = A\cos^2\tau + 2B\cos\tau\sin\tau + C\sin^2\tau, \tag{33}$$

and it follows that

$$E = \frac{A}{(AC - B^2)^2}, \quad F = \frac{B}{(AC - B^2)^2}, \quad G = \frac{C}{(AC - B^2)^2}.$$

But it is also true that Φ must be such a homogeneous expression in u, $\cos\tau$, $\sin\tau$. For by differentiating the left member of (32) for τ it follows that

$$\frac{d^3\Phi}{d\tau^3} + 4\frac{d\Phi}{d\tau} = 0,$$

and therefore

$$u_\tau^3\Phi_{uuu} + 3u_\tau^2\Phi_{uu\tau} + 3u_\tau(\Phi_{u\tau\tau} - u\Phi_{uu} + \Phi_u) + (\Phi_{\tau\tau\tau} - 3u\Phi_{u\tau} + 4\Phi_\tau) = 0.$$

This is an identity in x, y, τ and must also be an identity in τ, u, u_τ, since a set of values (x, y, τ) can be determined which gives τ, u, u_τ any arbitrarily assigned values. Hence by equating coefficients of powers of u_τ to zero, it can be readily seen that

$$\Phi = u^2K(\tau) + uL(\tau) + M(\tau),$$

where

$$\frac{dK}{d\tau} = 0, \quad \frac{d^2L}{d\tau^2} + L = 0, \quad \frac{d^3M}{d\tau^3} + 4M = 0,$$

from which it follows that K is a constant while L and M are any homogeneous expressions of the first and second degrees, respectively, in $\cos\tau$ and $\sin\tau$.

Hence the only possible length integrals for surfaces whose geodesics are to be represented by straight lines in the xy-plane, i. e., the plane of the parameters of the surface, are integrals of the form

$$f(x, y, \tau) = \frac{\sqrt{\Phi(u, \cos\tau, \sin\tau)}}{AC - B^2}.$$

In this expression Φ may be any homogeneous polynomial of the second order in the three variables $u = x\sin\tau - y\cos\tau$, $\cos\tau$, $\sin\tau$, and A, B, C are the coefficients of its expansion (33) in powers of $\cos\tau$ and $\sin\tau$.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

A GEOMETRICAL INTERPRETATION OF THE GENERALIZED LAW OF THE MEAN

By Charles N. Haskins

The geometrical interpretation of the ordinary Law of the Mean,

$$(1) \qquad f(b) - f(a) = (b - a) f'(X), \qquad a < X < b,$$

and of the function

$$(2) \qquad \phi(x) \equiv \frac{f(b) - f(a)}{b - a}(x - a) - [f(x) - f(a)],$$

by which its proof is made to depend on Rolle's theorem, are well-known and may be stated as follows. Consider the curve $y = f(x)$, the points $A: \{a, f(a)\}$, $B: \{b, f(b)\}$ upon it, and the chord AB. Then there exists a point $P: \{X, f(X)\}$ on the curve between A and B such that the tangent at P is parallel to the chord AB. The function $\phi(x)$ represents the distance from the variable point $\{x, f(x)\}$ of the curve to the chord, this distance being measured parallel to the axis of Y.

The importance of the generalized law of the mean, especially in the discussion of indeterminate forms, is well recognized,* but its simple geometric interpretation and that of the function by means of which its proof is made to depend on Rolle's theorem have not been emphasized as strongly as is desirable if the law is to be useful in elementary instruction in the calculus.

The Generalized Law of the Mean is

$$(3) \qquad \frac{f(b) - f(a)}{F(b) - F(a)} = \frac{f'(X)}{F'(X)}, \qquad a < X < b.$$

The function $\Phi(x)$ used in the proof is

$$(4) \qquad \Phi(x) \equiv \frac{f(b) - f(a)}{F(b) - F(a)}[F(x) - F(a)] - [f(x) - f(a)].$$

* Cf. Osgood, Annals of Mathematics, ser. 1, vol. 12, p. 65, 1899; and *Differential and Integral Calculus*, p. 230.

The interpretations are as follows. Consider the space curve, $y = f(x)$, $z = F(x)$, the points $A: \{a, f(a), F(a)\}$, $B: \{b, f(b), F(b)\}$ upon it, the chord AB, and the plane determined by the chord AB and the straight line through A parallel to the axis of X. Call this plane the *chordal plane*. Then as interpretation of the law (3) we have:

There exists a point $P: \{X, f(X), F(X)\}$ on the curve such that the tangent at P is parallel to the chordal plane.

And as interpretation of the function $\Phi(x)$ in (4):

The function $\Phi(x)$ represents the distance from the variable point $\{x, f(x), F(x)\}$ of the curve to the chordal plane, this distance being measured parallel to the axis of Y.

The proofs of these statements are at once obvious if it is noted that the equation of the chordal plane is

$$y = \frac{f(b) - f(a)}{F(b) - F(a)} [z - F(a)] + f(a).$$

When once this geometric representation of the law and of the auxiliary function $\Phi(x)$ has been obtained, the reason for the appearance of this function is clear, and the course of the proof is seen to be identical with that of the simpler law (1).

The more general forms of the Law of the Mean expressed by*

$$\begin{vmatrix} \phi'(T) & f'(T) & F'(T) \\ \phi(a) & f(a) & F(a) \\ \phi(b) & f(b) & F(b) \end{vmatrix} = 0, \qquad a < T < b;$$

and

$$\begin{vmatrix} \phi'(T_i) & f'(T_i) & F'(T_i) & 0 \\ \phi(a) & f(a) & F(a) & 1 \\ \phi(b) & f(b) & F(b) & 1 \\ \phi(c) & f(c) & F(c) & 1 \end{vmatrix} = 0, \qquad \begin{matrix} i = 1, 3, \\ a < T_1 < b < T_3 < c; \end{matrix}$$

* Cf. Jordan, *Cours d'Analyse*, vol. 1, p. 66.

$$\begin{vmatrix} \phi''(T_2) & f''(T_2) & F''(T_2) & 0 \\ \phi(a) & f(a) & F(a) & 1 \\ \phi(b) & f(b) & F(b) & 1 \\ \phi(c) & f(c) & F(c) & 1 \end{vmatrix} = 0, \qquad a < T_1 < T_2 < T_3 < c;$$

admit a similar interpretation by means of the space curve: $x = \phi(t)$, $y = f(t)$, $z = F(t)$. The first expresses the fact that if

$$A: \{\phi(a), f(a), F(a)\}, \qquad B: \{\phi(b), f(b), F(b)\}$$

are two points of a space curve, there is between A and B a point T on the curve where the tangent is parallel to the plane determined by A, B, and the origin. The second and third state that if A, B, C, are three non-collinear points of a space curve, there is a point T_1 on the curve between A and B, and a point T_3 between B and C, where the tangent to the curve is parallel to the plane $A\ B\ C$; and further there is a point T_2 between T_1 and T_3 where the principal normal is parallel to that plane.

UNIVERSITY OF ILLINOIS,
DECEMBER, 1907.

ON THE TORSION OF A CURVE

By Paul Saurel

The object of this note is to call attention to a new way of stating the definition of the torsion of a curve, and to show that the new definition yields immediately the familiar formulas for the torsion of a space curve, of a geodesic curve, and of any curve traced on a surface.

The torsion of a curve at a given point is, by definition, the limit of the ratio of the angle between the osculating planes at the given point and at a neighboring point, to the arc joining these points. And in order that the sign of the torsion as thus determined shall agree with the sign as determined by Frenet's formulas it is necessary that the angle between the two planes be considered positive or negative according as the angle between the binormal at the given point and the principal normal at the neighboring point is obtuse or acute.*

This definition can be transformed into the following: the torsion of a curve at a given point is the limit of the ratio of the angle between the osculating plane at the given point and the principal normal at the neighboring point, to the arc joining the two points. To justify this definition we observe that the angle between the osculating plane at the given point and the tangent line at the neighboring point is an infinitesimal of the second order when the arc joining the two points is of the first order.† This tangent may therefore be said to lie in the osculating plane at the given point and it may thus be considered as the intersection of the two osculating planes. The angle between the two osculating planes may accordingly be replaced by the angle between the osculating plane at the given point and the principal normal at the neighboring point.

In applying this definition we can, of course, replace the angle between the osculating plane at the given point and the principal normal at the neighboring point by its sine, and this, in turn, by the cosine of the angle between the binormal at the given point and the principal normal at the neighboring

*Cf. Kommerell und Kommerell, *Allgemeine Theorie der Raumkurven und Flächen*, vol. 1, p. 24.

†Cf. C. Jordan, *Cours d'analyse*, vol. 1, p. 465.

point. Moreover, if we remember the convention stated above concerning the sign of the angle between the two osculating planes, it is easy to see that this angle, with its proper sign, is equal to the negative of the cosine of the angle between the binormal at the first point and the principal normal at the second point.

It is worth noticing that our definition holds if the principal normal instead of being drawn as usual toward the centre of curvature be drawn in the opposite direction, provided the tangent, the principal normal and the binormal form a set of axes congruent with the axes of reference.

These results can of course be established by analysis. If we denote the arc of the curve by s, the radius of curvature by r, the radius of torsion by ρ, and the direction cosines of the tangent, the principal normal and the binormal by α, β, γ, l, m, n, λ, μ, ν, respectively, and if we suppose, moreover, that the positive direction of the principal normal passes through the centre of curvature and that the tangent, the principal normal, and the binormal form a system of axes congruent with the axes of reference, Frenet's formulas will take the form

$$\frac{d\alpha}{ds} = \frac{l}{r}, \qquad \frac{d\beta}{ds} = \frac{m}{r}, \qquad \frac{d\gamma}{ds} = \frac{n}{r}, \tag{1}$$

$$\frac{dl}{ds} = -\frac{\alpha}{r} - \frac{\lambda}{\rho}, \qquad \frac{dm}{ds} = -\frac{\beta}{r} - \frac{\mu}{\rho}, \qquad \frac{dn}{ds} = -\frac{\gamma}{r} - \frac{\nu}{\rho}, \tag{2}$$

$$\frac{d\lambda}{ds} = \frac{l}{\rho}, \qquad \frac{d\mu}{ds} = \frac{m}{\rho}, \qquad \frac{d\nu}{ds} = \frac{n}{\rho}. \tag{3}$$

The direction cosines of the binormal at any point being λ, μ, ν and those of the principal normal at an adjacent point being $l + \Delta l$, $m + \Delta m$, $n + \Delta n$, the cosine of the angle between the two lines is equal to the expression

$$\lambda(l + \Delta l) + \mu(m + \Delta m) + \nu(n + \Delta n), \tag{4}$$

or, more simply, to

$$\lambda\,\Delta l + \mu\,\Delta m + \nu\,\Delta n. \tag{5}$$

If in this expression we replace Δl, Δm, Δn by their values obtained from equations 2, we get

$$-\lambda\left(\frac{\alpha}{r} + \frac{\lambda}{\rho}\right)\Delta s - \mu\left(\frac{\beta}{r} + \frac{\lambda}{\rho}\right)\Delta s - \nu\left(\frac{\gamma}{r} + \frac{\nu}{\rho}\right)\Delta s + \cdots, \tag{6}$$

the omitted terms being of the second degree in Δs, and this in turn reduces to

$$-\frac{1}{\rho}\,\Delta s + \cdots. \tag{7}$$

Thus the limit of the fraction obtained by dividing the negative of this expression by the increment of the arc is the torsion.

If we suppose that the positive direction of the principal normal does not pass through the centre of curvature but that nevertheless the tangent, the principal normal and the binormal form a set of axes congruent with the axes of reference, we must replace $l, m, n, \lambda, \mu, \nu$ in equations 1, 2, 3 by $-l, -m, -n, -\lambda, -\mu, -\nu$. The new formulas lead again to expression 7.

Let us apply our definition to any curve in space. We shall suppose that the coordinates x, y, z of any point on the curve are given as functions of the arc s measured from some given point of the curve, and we shall use primes to indicate differentiations with respect to s. Then, as is well known, the direction cosines of the tangent at any point are equal to x', y', z', while those of the principal normal and of the binormal are respectively proportional to

$$x'', \qquad y'', \qquad z'', \tag{8}$$

$$y'z'' - y''z', \quad z'x'' - z''x', \quad x'y'' - x''y'. \tag{9}$$

In both cases the factor of proportionality is the reciprocal of $\sqrt{x''^2 + y''^2 + z''^2}$.

From 8 it follows that the direction cosines of the principal normal at a neighboring point are proportional to

$$x'' + x'''\Delta s, \quad y'' + y'''\Delta s, \quad z'' + z'''\Delta s, \tag{10}$$

in which Δs denotes the increment of s. The factor of proportionality is now approximately equal to the reciprocal of $\sqrt{x''^2 + y''^2 + z''^2}$. From 9 and 10 we can calculate the cosine of the angle between the binormal at the given point and the principal normal at the adjacent point. If we divide this by $-\Delta s$ and take the limit we obtain at once the well known formula for the torsion $1/\rho$ of a space curve:

$$\frac{1}{\rho} = -\frac{\begin{vmatrix} x' & y' & z' \\ x'' & y'' & z'' \\ x''' & y''' & z''' \end{vmatrix}}{x''^2 + y''^2 + z''^2}. \tag{11}$$

Let us next consider a geodesic curve on a given surface. Take the given point on the curve as the origin, the tangents to the lines of curvature through the point as axes of x and y, and the normal to the surface as axis of z. Denote, as usual, the first derivatives of z with respect to x and y by p and q, and the second derivatives by r, s, t. Then, because of our choice of axes, at the origin p, q and s are equal to zero, and at a neighboring point p and q are approximately equal to $rx'\Delta s$, $ty'\Delta s$ in which Δs denotes the arc between the two points and the primes indicate differentiation with respect to s.

The direction cosines of the tangent to the geodesic curve at the origin are

$$x', y', 0. \tag{12}$$

Since the principal normal at each point of a geodesic curve is, by definition, normal to the surface, its direction cosines are proportional to $-p$, $-q$, 1. At the origin the direction cosines of the principal normal to the geodesic line are therefore equal to

$$0, 0, 1, \tag{13}$$

and at an adjacent point of the geodesic line the direction cosines of the principal normal are approximately equal to

$$-rx'\Delta s, -ty'\Delta s, 1. \tag{14}$$

From 12 and 13 we find the direction cosines of the binormal to the geodesic line at the origin to be

$$y', -x', 0. \tag{15}$$

From 14 and 15 we can calculate the cosine of the angle between the binormal to the curve at the origin and the principal normal to the curve at an adjacent point. If we divide this by $-\Delta s$ we get the formula for the torsion $1/\rho_g$ of a geodesic curve

$$\frac{1}{\rho_g} = (r - t)x'y'. \tag{16}$$

If we denote the principal radii of curvature by R_1, R_2 and the angle between the geodesic line and one of the lines of curvature by θ, the equation assumes the more familiar form

$$\frac{1}{\rho_g} = \left(\frac{1}{R_1} - \frac{1}{R_2}\right)\cos\theta\sin\theta = \frac{1}{2}\left(\frac{1}{R_1} - \frac{1}{R_2}\right)\sin 2\theta. \tag{17}$$

Finally, let us consider any curve drawn on a surface. At the given point P of the curve draw the tangent geodesic line. The two curves will have in common not only the point P but also an adjacent point Q. Let us denote by ω the angle that the principal normal to the given curve makes with the normal to the surface. Then, if we remember that the normal to the surface is also the principal normal of the geodesic line, it is not hard to see that the angle between the osculating plane to the given curve at P and the principal normal to the given curve at Q is equal to the angle between the osculating plane to the geodesic curve at P and the principal normal to the geodesic curve at Q plus the increment of ω. If then we denote the torsion of the tangent geodesic curve by $1/\rho_g$, it follows that the torsion $1/\rho$ of the given curve is determined by the equation

$$\frac{1}{\rho} = \frac{1}{\rho_g} + \omega', \tag{18}$$

in which the prime indicates differentiation with respect to the arc s. The positive direction of ω is from the binormal to the principal normal of the given curve.

NEW YORK,
NOVEMBER, 1907.

ON THE SPHERICAL REPRESENTATION OF A SURFACE

By Paul Saurel.

The focal lines corresponding to a given point on a surface are two straight lines each of which is drawn through one of the principal centres of curvature, parallel to the line of curvature that corresponds to the other principal centre. And it is well known that all the normals to the surface drawn from the neighborhood of the given point intersect these two lines.* In this note I should like to indicate how this property of the focal lines enables us to state in geometric terms a very elegant demonstration, due to Stäckel,† of a fundamental relation in the theory of the spherical representation of a surface.

Let P and Q be neighboring points on a surface and let the projections of PQ on the axes be dx, dy, dz. If P' and Q' be the points in which the lines through the origin parallel to the normals at P and Q pierce the surface of the unit sphere whose centre is at the origin, and if X, Y, Z be the direction cosines of the normal at P, the projections of $P'Q'$ on the axes will be dX, dY, dZ.

From the point F_1, in which the normal through Q cuts the focal line parallel to the first line of curvature through P, draw a line parallel to the normal at P; this line will cut the first line of curvature in a point P_1. Since the focal line on which F_1 lies passes through the second centre of curvature, the distance F_1P_1 will be equal to the second radius of curvature R_2. Moreover, the line QP_1 will be parallel to $Q'P'$ and its projections on the axes will be equal, in magnitude and in sign, to R_2dX, R_2dY, R_2dZ. It follows that the projections of PP_1, the geometric sum of PQ and QP_1, are equal to

$$dx + R_2dX, \quad dy + R_2dY, \quad dz + R_2dZ. \tag{1}$$

* Cf. G. Scheffers, *Einführung in die Theorie der Flächen*, p. 166.

† P. Stäckel, *Bulletin des sciences mathématiques*, 2nd series, vol. 27, p. 139; 1903.

In like manner, through the point F_2 in which the normal at Q cuts the second focal line, draw a line parallel to the normal at P; this line will cut the second line of curvature in a point P_2. If we denote the first radius of curvature by R_1, the projections of QP_2 will be R_1dX, R_1dY, R_1dZ, and those of PP_2

$$dx + R_1dX, \quad dy + R_1dY, \quad dz + R_1dZ. \tag{2}$$

If we remember that the lines of curvature are at right angles to each other, we get at once from (1) and (2) the equation

$$\sum (dx + R_2dX)(dx + R_1dX) = 0, \tag{3}$$

and from this we obtain without difficulty the fundamental relation

$$\frac{1}{R_1R_2}\sum dx^2 + \left(\frac{1}{R_1} + \frac{1}{R_2}\right)\sum dx\,dX + \sum dX^2 = 0. \tag{4}$$

New York,
May, 1907.

THE ABSOLUTE MINIMUM IN THE PROBLEM OF THE SURFACE OF REVOLUTION OF MINIMUM AREA

By Mary Emily Sinclair

Introduction. It is well known that the problem of finding a relative minimum of the integral

$$I = \int y\sqrt{x'^2 + y'^2}\, dt$$

with respect to all ordinary * curves which can be drawn in the region

$$R:\ y \geqq 0$$

from a given point A_0 to another given point A_1, has either one or two solutions, according to the position of the two points A_0 and A_1. When the two points are sufficiently far apart there is only one solution, the Goldschmidt discontinuous solution,† consisting of the ordinates of A_0 and A_1 together with the segment of the x-axis between them. If A_0 and A_1 are near to each other there are two solutions, the discontinuous solution just described, and, besides, an arc of a catenary with the x-axis as directrix.

These known results would immediately furnish also the solution of the problem of the *absolute* minimum of the above integral, *if we knew a priori the existence of an absolute minimum.* But such an existence proof ‡ has never been given, and the problem of the absolute minimum of the integral is, as far as I know, still unsolved. To fill this gap in the theory of the surface of revolution of minimum area is the object of the following pages.

We make use of the following lemma, due essentially to Todhunter: §

---

* In the terminology of Bolza: *Lectures on the Calculus of Variations*, p. 117.

† Goldschmidt, *Göttingen Prize Essay*, 1831. Todhunter, *Researches in the Calculus of Variations*, p. 60. Bolza, loc. cit., p. 153.

‡ The assumptions under which Hilbert's existence theorem has been proved by Bolza (loc. cit. p. 247) are not satisfied for the region $R:\ y \geqq 0$.

§ Todhunter: loc. cit., p. 60, §§64, 65.

Lemma: *If M_0A_0 is the ordinate of A_0, and if A_0L_0 is the arc of any ordinary curve not identical with A_0M_0, and such that arc $A_0L_0 = |A_0M_0|$, then,*

$$I(A_0L_0) > I(A_0M_0)^*$$

Let P and P' be any two points on A_0M_0 and A_0L_0 respectively, such that

$$|A_0P| = \text{arc } A_0P',$$

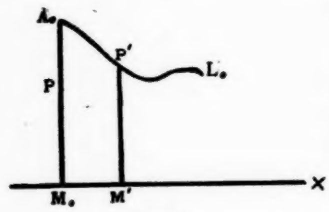


Fig. 1.

and let $M'P'$ be the ordinate of P'. When A_0P' coincides with A_0P, we have $|M'P'| = |M_0P|$. When A_0P' is different from A_0P, it follows that $|M'P'| > |M_0P|$ since a normal is the shortest distance from a point to a line.

Since L_0 does not coincide with M_0, there must be points P' of A_0L_0 for which A_0P' does not coincide with A_0P, the corresponding segment of A_0M_0. For such points

$$|M'P'|\,ds > |M_0P|\,ds,$$

and hence,
$$\int |M'P'|\,ds > |M_0P|\,ds,$$

or
$$I(A_0L_0) > I(A_0M_0).$$

This leads to the following theorem: *The Goldschmidt solution consisting of the path $A_0M_0M_1A_1$, furnishes a value of I less than that furnished by any other ordinary curve $\mathfrak{C}$ in $R(y \geqq 0)$ joining A_0 and A_1, whose length is equal to or greater than $|A_0M_0| + |A_1M_1|$.*

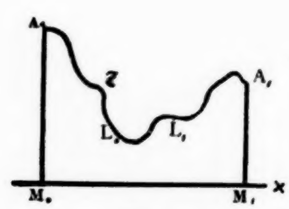


Fig. 2.

For let the arc A_0L_0 of $\mathfrak{C}$ equal $|A_0M_0|$, and let the arc A_1L_1 of $\mathfrak{C}$ equal $|A_1M_1|$. When A_0L_0 coincides with A_0M_0, and A_1L_1 with A_1M_1,

---

* $I(AB) = I(BA)$.

then $$I_{\mathfrak{C}}(A_0L_0) = I(A_0M_0),$$

and $$I_{\mathfrak{C}}(A_1L_1) = I(A_1M_1).$$

But $I(M_0M_1) = 0$ and $I_{\mathfrak{C}}(L_0L_1) > 0$, since $\mathfrak{C}$ cannot coincide throughout with $A_0M_0\ M_1A_1$. Hence, $I_{\mathfrak{C}}(A_0A_1) > I(A_0M_0M_1A_1)$.

When A_0L_0 and A_1L_1 do not both coincide with A_0M_0 and A_1M_1, respectively, then, by the lemma,

$$I_{\mathfrak{C}}(A_0L_0) + I_{\mathfrak{C}}(A_1L_1) > I(A_0M_0) + I(A_1M_1).$$

Hence

$$I_{\mathfrak{C}}(A_0A_1) \geqq I_{\mathfrak{C}}(A_0L_0) + I_{\mathfrak{C}}(A_1L_1) > I(A_0M_0M_1A_1).$$

Hence the theorem is always true.

COR. $I_{\mathfrak{C}}(AB) > I(A_0M_0M_1A_1)$ *if* $\mathfrak{C}$ *is any rectifiable curve whose length is greater than* $|A_0M_0| + |A_1M_1|$. For, let arc $A_0L_0 = |A_0M_0|$ and arc $A_1L_1 = |A_1M_1|$, let N_0 and N_1 be two points of $\mathfrak{C}$ in the order $L_0N_0N_1L_1$. Consider any partition π, and a polygon P_π inscribed in A_0N_0. Then, if the partition is sufficiently fine, the length of P_π is greater than the arc A_0L_0. According to the Todhunter lemma,

$$I_{P_\pi}(A_0N_0) \geqq I(A_0M_0),^*$$

and, passing to the limit,†

$$I^*_{\mathfrak{C}}(A_0N_0) \geqq I(A_0M_0).$$

Similarly, $$I^*_{\mathfrak{C}}(A_1N_1) \geqq I(A_1M_1).$$

Moreover, $$I^*_{\mathfrak{C}}(N_0N_1) > 0.$$

Hence, $$I^*_{\mathfrak{C}}(A_0A_1) > I(A_0M_0M_1A_1).$$

We may now apply the above results to our problem. We consider two cases.

Case I. $|A_0A_1| \leqq y_0 + y_1$. Then the length of every ordinary curve joining A_0 and A_1 is equal to or greater than $y_0 + y_1$. By the above theorem,

---

* We use here the notation and definition of Bolza, loc. cit., p. 157, b).

† The integral $I = \int y\sqrt{x'^2 + y'^2}\,dt$ has not in general a well defined meaning unless the curve over which it is taken is represented by functions $x = x(t)$, $y = y(t)$ which have derivatives. The integrals $I_{P_\pi}(A_0N_0)$ have a limit, however, whenever the polygons over which they are taken are inscribed in a rectifiable curve. This limit is a generalization of the value of I denoted by I^*. See Bolza, loc. cit., p. 159.

the Goldschmidt solution furnishes therefore an absolute minimum for all ordinary curves joining A_0 and A_1 in the region R.

Case II. $|A_0A_1| < y_0 + y_1$. We construct the ellipse whose foci are A_0 and A_1 and whose points P satisfy the equation

$$|A_0P| + |A_1P| = y_0 + y_1,$$

and denote by R_0 the closed region consisting of the ellipse and its interior.

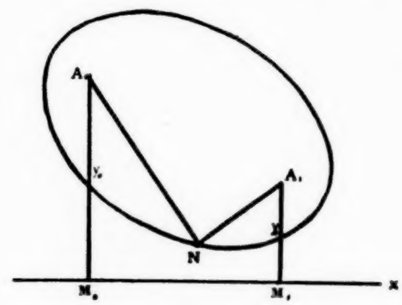


FIG. 3.

Then every point in the region R_0^- is such that its ordinate y is greater than zero. For, if not, let N be one point of the ellipse, lying on the x-axis. Then $|A_0N| > y_0$ unless N coincides with M_0, and $|A_1N| > y_1$ unless N coincides with M_1, and therefore $|A_0N| + |A_1N| > y_0 + y_1$, since N cannot coincide with both M_0 and M_1. But this contradicts the hypothesis,

$$|A_0N| + |A_1N| = y_0 + y_1.$$

For the region R_0, the hypotheses of the Hilbert existence theorem* are fulfilled, since

A) $F(x_1y, kx', ky') = y\sqrt{kx'^2 + k^2y'^2} = kF(xyx'y')$,

B) $F(xy \cos\gamma \sin\gamma) = y > 0$,

C) $F_1(xy \cos\gamma \sin\gamma) = y > 0$,

D) R_0 is convex and A_0 and A_1 are interior points.

Hence we may apply the Hilbert existence theorem to the region R_0^-, obtaining the result that among all rectifiable curves in R_0 joining A_0 and A_1 there

* Bolza, loc. cit., §43.

exists at least one curve $\mathfrak{C}$ furnishing an absolute minimum for the integral I. This curve is either an extremal $\mathfrak{E}$ of class C', i. e., a catenary with x-axis as directrix, or consists of portions of extremals separated by points of the boundary R_0. In the latter case, its length would be greater than $y_0 + y_1$, for it could not consist merely of two straight lines from A_0 and A_1 to a single point P of the boundary, since these straight lines could not both be extremals. By the theorem proved above, such a curve would give a value of I greater than that given by the Goldschmidt solution.

Among all ordinary curves not wholly in R_0, the Goldschmidt solution furnishes an absolute minimum, since every such curve contains at least one point of the ellipse and its length is greater than $y_0 + y_1$. We have therefore the theorem:

Among all ordinary curves in R, $(y \geqq 0)$, joining A_0 and A_1, there exists a curve which furnishes for the integral $I = \int y ds$ an absolute minimum, and this curve is either a catenary through A_0 and A_1, not containing the conjugate point to A_0, if such exist, or it is the Goldschmidt discontinuous solution, $A_0M_0\ M_1A_1$.

Making use of the results of MacNeish* for relative minima, we now conclude:

1) *The catenary solution furnishes the absolute minimum if A_1 lies within the curve F defined for A_0 $(x_0\ y_0)$ by the equations*

$$F) \qquad \begin{aligned} x - x_0 &= a(t - t_0), \\ y = a \operatorname{ch} t, \ & y_0 = a \operatorname{ch} t_0, \\ 2t_0 + 1 + e^{2t_0} &= 2t - 1 - e^{-2t}. \end{aligned}$$

2) *Both solutions furnish the absolute minimum if A_1 lies on F, the values of I then being equal for the two solutions.*

3) *The Goldschmidt solution furnishes the absolute minimum if A_1 lies without F.*

OBERLIN COLLEGE,
OBERLIN, OHIO.

* MacNeish, ANNALS OF MATHEMATICS, vol. 7 (1906), p. 80.

NOTE ON THE ROOTS OF BESSEL FUNCTIONS

By Charles N. Moore

The earliest proof of the fact that the Bessel function of order zero, $J_0(x)$, has an infinite number of roots was given by Bessel in the year 1824* and was obtained directly from the integral definition of $J_0(x)$:

$$J_0(x) = \frac{2}{\pi}\int_0^1 \frac{\cos xt}{\sqrt{1-t^2}}\,dt. \tag{1}$$

In the course of his proof Bessel showed that $J_0(x)$ has at least one root in each of the intervals

$$(m+\tfrac{1}{2})\pi < x < (m+1)\pi \qquad (m = 0, 1, 2, \cdots) \tag{2}$$

and no positive roots outside of these intervals.

The object of this paper is to establish the fact that $J_0(x)$ has only one root in each of the intervals (2); this will be shown by applying methods analogous to those of Bessel to the integral definition of $J_1(x)$:†

$$J_1(x) = \frac{2x}{\pi}\int_0^1 \sqrt{1-t^2}\cos xt\,dt. \tag{3}$$

Although the result itself is not new it seems worth while, from a historical point of view at least, to give a proof of it which is analogous to Bessel's proof of the less specific fact.

The outline of the method is as follows: It is a well known fact that

$$\frac{dJ_0(x)}{dx} = -J_1(x).\ddagger$$

* Cf. *Berliner Abhandlungen* (1824), *Mathematische Klasse*, p. 39. For a reproduction of Bessel's proof see Gray and Mathews, *Treatise on Bessel Functions*, p. 44.

† This fact can also be obtained directly from the expression (1) but the proof in that case is somewhat more complicated.

‡ Cf. Gray and Mathews, *Treatise on Bessel Functions*, p. 13.

We shall prove that

$$\operatorname{sgn} J_1(x) = (-1)^m \qquad (m + \tfrac{1}{2})\pi \leqq x \leqq (m+1)\pi,$$

and hence that for this interval

$$\operatorname{sgn} \frac{dJ_0(x)}{dx} = (-1)^{m+1}.$$

Then according as m is odd or even $J_0(x)$ continually increases or continually decreases throughout the interval, and therefore in either case it cannot have more than one root in that interval. Combining this with the fact proved by Bessel, namely that $J_0(x)$ has at least one root in the interval, it follows that it has just one root in every such interval.

We consider values of x that lie in the interval

$$(m + \tfrac{1}{2})\pi \leqq x \leqq (m+1)\pi$$

and we set

$$x = \frac{2m + 1 + m'}{2}\pi \qquad (0 \leqq m' \leqq 1). \tag{4}$$

If now we make the change of variable

$$v = (2m + 1 + m')t$$

and for brevity put

$$\mu = 2m + 1 + m', \tag{5}$$

equation (3) reduces with the help of (4) to

$$J_1(x) = \frac{2x}{\pi}\int_0^{2m+1+m'} \frac{1}{\mu^2}\cos\frac{\pi v}{2}\sqrt{\mu^2 - v^2}\,dv. \tag{6}$$

If then we set

$$u_n = \int_{2n}^{2n+2} \cos\frac{\pi v}{2}\sqrt{\mu^2 - v^2}\,dv \qquad (n = 0, 1, 2, \cdots (m-1)), \tag{7}$$

and

$$I = \int_{2m}^{2m+1+m'} \cos\frac{\pi v}{2}\sqrt{\mu^2 - v^2}\,dv, \tag{8}$$

we may write (6) in the form

(9) $$J_1(x) = \frac{2x}{\pi\mu^2}\left[\sum_{n=0}^{n=m-1} u_n + I\right].$$

Since the quantity $\sqrt{\mu^2 - v^2}$, which appears as a factor in the integrand of the integral on the right-hand side of (7), continually decreases with increasing v as long as v lies in the interval $0 \leqq v \leqq \mu$ it follows that this quantity is less in absolute value in the second half of the interval of integration of the integral in question than it is in the first half. Hence the integral has the same sign that the integrand has in the first half of the interval, i. e.,

$$\operatorname{sgn} u_n = \operatorname{sgn} \cos n\pi = (-1)^n.$$

We see thus that the finite sum

$$\sum_{n=0}^{n=m-1} u_n$$

has its terms alternating in sign. We will show next that these terms are continually increasing in absolute value.

If we set

$$v' = v + 2,$$

we get for u_{n-1}

$$u_{n-1} = -\int_{2n}^{2n+2} \cos\frac{\pi v'}{2}\sqrt{\mu^2 - (v'-2)^2}\, dv'.$$

Consequently

$$u_n + u_{n-1} = \int_{2n}^{2n+2} \cos\frac{\pi v}{2}\left[\sqrt{\mu^2 - v^2} - \sqrt{\mu^2 - (v-2)^2}\right] dv.$$

The quantity in brackets in the integrand of the last integral is obviously negative throughout the interval of integration; moreover it decreases with increasing v since its derivative with regard to v is

$$-\frac{v}{\sqrt{\mu^2 - v^2}} + \frac{v-2}{\sqrt{\mu^2 - (v-2)^2}},$$

which is a negative quantity. It therefore increases in absolute value with increasing v and hence the integral in question has the sign of the integrand in

the second half of the interval, i. e.,

$$\operatorname{sgn}[u_n + u_{n-1}] = -\operatorname{sgn}\cos(n+1)\pi = (-1)^n = \operatorname{sgn} u_n,$$

and this shows, since u_n and u_{n-1} have opposite signs, that u_n is greater in absolute value than u_{n-1}.

We now wish to show that the quantity I, defined by (8), has the sign $(-1)^m$ and is greater in absolute value than u_{m-1}. By breaking up the integral in (8) we get

$$(10) \qquad I = \int_{2m}^{2m+1-m'} \cos\frac{\pi v}{2}\sqrt{\mu^2 - v^2}\,dv + \int_{2m+1-m'}^{2m+1+m'} \cos\frac{\pi v}{2}\sqrt{\mu^2 - v^2}\,dv.$$

The first integral on the right-hand side of (10) has the sign $(-1)^m$ since its integrand has that sign throughout the interval of integration; the second integral has the sign $(-1)^m$ since its integrand has that sign in the first half of the interval of integration and the quantity $\sqrt{\mu^2 - v^2}$ is greater in absolute value there than it is in the second half of the interval of integration. Hence I has the sign $(-1)^m$. If we can show further that the sum of I and u_{m-1} has the sign $(-1)^m$ it will follow that I is greater in absolute value than u_{m-1}, since the latter has the sign $(-1)^{m-1}$.

When we make the change of variable

$$v = v' + 2,$$

we get for I

$$I = -\int_{2m-2}^{2m-1+m'} \cos\frac{\pi v'}{2}\sqrt{\mu^2 - (v'+2)^2}\,dv',$$

so that we have

$$(11) \qquad u_{m-1} + I = \int_{2m-2}^{2m-1+m'} \cos\frac{\pi v}{2}\left[\sqrt{\mu^2 - v^2} - \sqrt{\mu^2 - (v+2)^2}\right]dv$$

$$+ \int_{2m-1+m'}^{2m} \cos\frac{\pi v}{2}\sqrt{\mu^2 - v^2}\,dv.$$

If, in the last integral, we set

$$v = 4m - 2 - v'$$

it becomes

$$-\int_{2m-2}^{2m-1-m'} \cos\frac{\pi v'}{2}\sqrt{\mu^2 - (4m-2-v')^2}\,dv',$$

so that we may write (11) in the form

(12) $u_{m-1} + I$

$$= \int_{2m-2}^{2m-1-m'} \cos\frac{\pi v}{2}\left[\sqrt{\mu^2 - v^2} - \sqrt{\mu^2 - (v+2)^2} - \sqrt{\mu^2 - (4m-2-v)^2}\right] dv$$

$$+ \int_{2m-1-m'}^{2m-1+m'} \cos\frac{\pi v}{2}\left[\sqrt{\mu^2 - v^2} - \sqrt{\mu^2 - (v+2)^2}\right] dv.$$

The expression in brackets in the integrand of the second integral on the right-hand side of (12) is obviously positive; moreover it increases with increasing v since its derivative

$$-\frac{v}{\sqrt{\mu^2 - v^2}} + \frac{v+2}{\sqrt{\mu^2 - (v+2)^2}}$$

is a positive quantity. Hence the integral in question has the sign of the integrand in the second half of the interval of integration, that is $(-1)^m$.

Let us write

(13) $$\psi(v) = \sqrt{\mu^2 - v^2} - \sqrt{\mu^2 - (v+2)^2} - \sqrt{\mu^2 - (4m-2-v)^2}.$$

On substituting the value $(2m-2)$ for v we get

(14) $$\psi(2m-2) = \sqrt{\mu^2 - (2m-2)^2} - 2\sqrt{\mu^2 - 4m^2}.$$

Since in view of (5) μ is greater than or equal to $2m+1$, we have

$$3\mu^2 \geqq 12m^2 + 12m + 3 > 12m^2 + 8m - 4,$$

or, adding $(\mu^2 - 16m^2)$ to both sides of the inequality,

$$4\mu^2 - 16m^2 > \mu^2 - 4m^2 + 8m - 4 = \mu^2 - (2m-2)^2.$$

This gives, on extracting square roots,

$$2\sqrt{\mu^2 - 4m^2} > \sqrt{\mu^2 - (2m-2)^2},$$

which, in combination with (14), shows that $\psi(2m-2)$ is a negative quantity.

Furthermore $\psi(2m-1)$ is either negative or zero, for we have from (13)

$$\psi(2m-1) = \sqrt{\mu^2-(2m-1)^2} - \sqrt{\mu^2-(2m+1)^2} - \sqrt{\mu^2-(2m-1)^2}$$
$$= -\sqrt{\mu^2-(2m+1)^2} \leqq 0.$$

In order to determine the sign of $\psi(v)$ for values of v lying between $(2m-2)$ and $(2m-1)$ we examine its second derivative. We have

$$\psi''(v) = -\frac{1}{\sqrt{\mu^2-v^2}} - \frac{v^2}{(\mu^2-v^2)^{\frac{3}{2}}} + \frac{1}{\sqrt{\mu^2-(v+2)^2}} \tag{15}$$
$$+ \frac{(v+2)^2}{(\mu^2-(v+2)^2)^{\frac{3}{2}}} + \frac{1}{\sqrt{\mu^2-(4m-2-v)^2}} + \frac{(4m-2-v)^2}{(\mu^2-(4m-2-v)^2)^{\frac{3}{2}}}.$$

The function $\psi''(v)$ is obviously positive for all values of v lying in the interval

$$2m-2 \leqq v \leqq 2m-1 \tag{16}$$

since the first term on the right-hand side of (15) is less in absolute value than the third and the second less in absolute value than the fourth throughout that interval. Hence $\psi'(v)$ is continually increasing throughout the interval (16).

We have then three possibilities for $\psi'(v)$. It may be negative throughout the interval (16), or it may be first negative and then positive, or it may be positive throughout the interval. In none of these cases can $\psi(v)$, which is negative at the beginning of the interval (16) and negative or zero at the end of it, become positive for any value of v lying in the interval. Hence we have

$$\operatorname{sgn} \psi(v) = -1 \qquad (2m-2 \leqq v \leqq 2m-1).$$

It follows, then, that the integrand of the first integral on the right-hand side of (12) has the sign $(-1)^m$ throughout the interval of integration, and therefore the integral has that sign. The whole right-hand side of (12) has then the sign $(-1)^m$. Consequently the left-hand side has that sign also and hence

$$|I| \geqq |u_{m-1}|.$$

Thus we see that the set of quantities

$$u_1, u_2, \cdots u_{m-1}, I$$

alternate in sign and continually increase in absolute value. Their sum, therefore, has the sign of I, for according as the number of terms is odd or even we can write that sum

$$u_1 + (u_2 + u_3) + \cdots + (u_{m-1} + I)$$

or

$$(u_1 + u_2) + (u_3 + u_4) + \cdots + (u_{m-1} + I).$$

We have then, in view of (9),

$$\operatorname{sgn} J_1(x) = (-1)^m \qquad (m + \tfrac{1}{2})\pi \leqq x \leqq (m+1)\pi,$$

and hence, as we have pointed out above, our theorem is proved.

Harvard University,
Cambridge, Mass.

A SMOOTH CLOSED CURVE COMPOSED OF RECTILINEAR SEGMENTS WITH VERTEX POINTS WHICH ARE NOWHERE DENSE *

By E. R. Hedrick

It has occurred to the writer repeatedly that examples of an interesting nature result from simple integration of known continuous functions whose behavior is remarkable.† An instance of this kind is the example presented in this note.

A curve

$$(1) \qquad \begin{cases} x = f(t), \\ y = \phi(t), \end{cases} \qquad (t_0 \leqq t \leqq t_1)$$

is smooth, according to Professor Osgood ‡ if it has a continuously turning tangent at every point, that is, if $f(t)$ and $\phi(t)$ possess continuous first derivatives which satisfy the relation $f'(t)^2 + \phi'(t)^2 > 0$ for every value of t considered.

It is remarkable that there exist curves formed by joining together straight line segments, which are smooth in this sense. Such a curve evidently cannot have a corner, for either $f'(t)$ or $\phi'(t)$ would fail to exist at such a point. In order to grasp clearly the example which follows it is desirable to repeat here in some detail a well-known example.§

Let us remove from the interval $0 \leqq x \leqq 1$ first the interior points of the middle thirds, i, e., the points $1/3 < x < 2/3$; of the two segments of

---

* Read at the summer meeting of the American Mathematical Society at Cornell University, Sept. 3, 1907.

† Another instance is given in a paper entitled "On a Function which occurs in the Law of the Mean," Annals of Mathematics, ser. 2, vol. 7, p. 177 (1906).

‡ See Osgood, *Lehrbuch der Functionentheorie*, Leipzig, Teubner, 1906, vol. 1, p. 122.

§ See, e. g., Osgood, l. c., p. 164 and 166; Lebesgue, *Leçons sur l'integration*, Paris, Gauthier-Villars, 1904, p. 13 and p. 26.

length 1/3 which remain let us remove the points interior to their middle third, i. e., the points $1/9 < x < 2/9$ and the points $7/9 < x < 8/9$; of the four segments of length 1/9 which remain let us remove the points interior to their middle thirds; and so on, the operation being carried on thus indefinitely. The points which remain after the whole operation is completed consist of the end-points of the intervals whose interiors have been removed, together with the limiting points of these end-points; we shall call this whole assemblage of points (E). The assemblage (E) is perfect and is nowhere dense in the interval $0 \leqq x \leqq 1$.

Let us now define a function $f(x)$ as follows: at a point $x = t$ which does not belong to (E), $f(t)$ shall be equal to the value of x at the middle point of the interval to which t belongs; at the points of (E), $f(x)$ is defined by the requirement that $f(x)$ shall be continuous, i. e., $\lim_{x=a} f(x) = f(a)$. This definition involves no contradiction; the function $f(x)$ defined is continuous in the interval $0 \leqq x \leqq 1$, it possesses a derivative and that derivative is zero except at the points of (E).* The figure is easy to construct; it consists of segments of straight lines parallel to the axis of x and of the same length as the intervals removed above, together with the limiting points of the end-points of these segments.†

Let us now form the integral of this function:

$$\phi(x) = \int_0^x f(x)dx; \tag{2}$$

the function $\phi(x)$ is defined for every value of x in the interval $0 \leqq x \leqq 1$, since $f(x)$ is everywhere continuous; moreover, the derivative of $\phi(x)$ exists at every point $0 \leqq x \leqq 1$, the derivatives at the points $x = 0$ and $x = 1$ being, of course, one-sided; and $\phi'(x) = f(x)$.

The curve

$$y = \phi(x) \tag{3}$$

is continuous and its derivative $\phi'(x)$ $[=f(x)]$ exists and is continuous at every point of the interval $0 \leqq x \leqq 1$. Since $f(x)$ is a constant in each of

* The results quoted at this point and below are proved, for example, in the books just mentioned.

† It is important to visualize this figure; it is drawn, e. g., by Osgood, l. c.

the intervals removed to form (E), $\phi(x)$ is a linear function and the curve (3) is a straight line throughout each of those intervals. It is therefore clear that (3) represents a curve which is *smooth* in the sense above and which is composed of straight line segments except for the end-points of these line segments and the limiting points of these end-points; and *these exceptional points (vertices) are nowhere dense*, i. e., on any arc of the curve (3) there is at least one partial arc which is a straight line. Even at these exceptional points the curve is *smooth*. Nevertheless the whole curve (3) is not a straight line, for $\phi'(0) = f(0) = 0$ and $\phi'(1) = f(1) = 1$.

A combination of eight arcs, symmetrically arranged, similar to that defined by (3) evidently gives rise to a curve which is *closed*, *smooth* in Osgood's sense, and which is composed of rectilinear segments with the exception of a set of points which is not dense on any arc of the curve.

It is interesting to observe that (3) consists of the chords of the parabola $y = \frac{1}{2}x^2$ which connect those points of the parabola whose abscissae are the end-points of the intervals mentioned above, together with the limiting points of the end-points of these chords.

For, let $x = e$ be a point of (E); the area

$$\int_0^e f(x)dx = \phi(e)$$

may be thought of as made up of the areas under an arbitrary (but fixed) number of these rectilinear segments of $f(x)$ which lie above the interval $0 \leqq x \leqq e$, together with a residual area. The former area is obviously the same as the area under the corresponding portions of the curve $y = x$; the residual area can be made arbitrarily small, since the content of (E) is *nil*. Hence the total area is the same as that under $y = x$, or

$$\phi(e) = \int_0^e x\,dx = \tfrac{1}{2}e^2.$$

The same fact results at once from the consideration that the points of (E) are "of measure zero," whence they may be neglected in computing the integral.*

* To show this, it is necessary to consider that the assemblage (E) is "of measure zero;" hence any or all of its points may be neglected in determining the integral. See, e. g., Lebesgue, l. c.

It follows that the closed curve mentioned is built up upon eight arcs of parabolas which fit together smoothly at their end-points. It would be equally simple, following this construction, to build up a closed curve of the same type composed of chords of a circle, together with the limiting points of the end-points of these chords. Had we proceeded in this manner, however, the method of reaching the result would have been disguised, and it would have been necessary to give an independent proof of the statements which we have been able to throw back upon the well-known results for the assemblage (E) and the function $f(x)$.

COLUMBIA, MO.,
JANUARY, 1908.

EVALUATION OF THE PROBABILITY INTEGRAL

By Frank Gilman

A simple and convenient formula for calculating the numerical value of the definite integral

$$\int_0^x e^{-x^2}\,dx$$

has long seemed to me a desideratum. The three series commonly used for computing this integral are not convenient in practice, except when the argument x is either very small or very large. It is believed that the formulas here given will be found equally convenient for any value of x from 0 to 3, and that they will, in general, give the integral correct to five figures. It is true that some of the old tables give this integral to eleven, and some even to fourteen decimals; but so great a degree of accuracy appears to be unnecessary. Besides, the method here given can be extended so as to secure any degree of accuracy desired.

The integral

$$\int_0^x e^{-x^2}\,dx$$

is of great importance in many branches of applied mathematics. It is used in discussing the effect of refraction,* in investigating the secular cooling of the earth † and in discussions in regard to the conduction of heat, as found in Fourier's work, *Théorie analytique de la chaleur.*

---

* See Chauvenet's *Astronomy*, vol. 1.

† See Thompson and Tait's *Natural Philosophy*, vol. 1, p. 717.

It is of the greatest importance in the theory of probability, or probability of errors, and on this account, when written in the form

$$\int_x^\infty e^{-x^2}\,dx = \operatorname{erf} x$$

it is sometimes called the error function of x, while

$$\int_0^x e^{-x^2}\,dx = \operatorname{erfc} x$$

is called the error function complement of x. The former of these is designated by the notation erf x, and the latter by erfc x. The two are so related that

$$\operatorname{erf} x + \operatorname{erfc} x = \int_0^\infty e^{-x^2}\,dx,$$

or

$$\operatorname{erf} x + \operatorname{erfc} x = \tfrac{1}{2}\sqrt{\pi}.$$

We shall give formulas for computing directly erfc x, since, this being known, erf x can be immediately deduced.

Besides the applications of erf x above mentioned, it is of great use in the integration of many differential equations (Kramp says an infinite number) and Glaisher in the *Philosophical Magazine* for 1871,* has given a list of some of the definite integrals which can be expressed in terms of it.

Glaisher says that from its uses in physics, erf x may fairly claim at present to rank in importance next to the trigonometrical and logarithmic functions.

The following is the formula proposed for computing erfc x for values of x from 0 to 1:

$$\int_0^x e^{-x^2}\,dx = A + ax + bx^3 + cx^5. \tag{1}$$

When $x > 1$ we may write

$$\int_0^x e^{-x^2}\,dx = A + \frac{a}{x} + \frac{b}{x^3} + \frac{c}{x^5}. \tag{2}$$

---

* Vol. 42, pp. 294, 421.

The following are the values of the constants A, a, b, and c, corresponding to assigned limits of x.

Limits of x	A	Log a	Log b	Log c
0 to 0.40	0	0	9.5221346n	8.9674202
0.40 " .55	0.000225	9.9995025	9.5148674n	8.9090072
.55 " .70	.002127	9.9966907	9.4956431n	8.8260698
.70 " .85	.009435	9.9882445	9.4598519n	8.7203776
.85 " 1.00	.024000	9.9739908	9.4159907n	8.6187244
1.00 " 1.15	.960956	8.8864253n	9.2954356n	8.7802668
1.15 " 1.30	.871946	9.0553760	9.5600979n	9.0974426
1.30 " 1.60	.809877	9.4213859	9.7244149n	9.3195134
1.60 " 2.00	.831292	9.2883208	9.5942662n	8.9499540
2.00 " 2.40	.883119	7.5394525	8.9915539	9.6821021n
2.40 " 3.00	.893330	8.6048101n	9.3971250	9.8547853n

The following was the method of calculating these quantities. Assume

$$a' + b'x^2 + c'x^4 = e^{-x^2}.$$

Substitute in this equation certain values of x, and from the condition equations thus obtained, form the normal equations, according to the method of least squares. The solution of these equations will give the values of a', b', and c'. Then by integration we obtain

$$\int_0^x e^{-x^2}dx = A + a'x + \frac{b'}{3}x^3 + \frac{c'}{5}x^5,$$

in which A is a constant, $a = a'$, $b = \frac{1}{3}b'$, $c = \frac{1}{5}c'$.

When $x > 1$ we write

$$\frac{a'}{x^2} + \frac{b'}{x^4} + \frac{c'}{x^6} = e^{-x^2}$$

and proceed in the same way, remembering that $a = -a'$, $b = -\frac{1}{3}b'$, $c = -\frac{1}{5}c'$.

In order to explain more fully the method in detail, let us find the coeffi-

cients A, a, b, and c corresponding to the limits of x: 1.60 to 2.00. Substitute in the equation

$$\frac{a'}{x^2} + \frac{b'}{x^4} + \frac{c'}{x^6} = e^{-x^2}$$

the following successive values of x: 1.6, 1.65, 1.7, 1.75, 1.8, 1.85, 1.9, 1.95, 2.00 and write the 9 equations of condition:

$$\begin{aligned}
0.39062496a' + 0.15258786b' + 0.05960463c' &= 0.07730473,\\
.36730950a' + .13491630b' + .04955605c' &= .06571026,\\
.34602080a' + .11973039b' + .04142920c' &= .05557621,\\
.32653068a' + .10662229b' + .03481545c &= .04677061,\\
.30864200a' + .09525987b' + .02940120c' &= .03916391,\\
.29218412a' + .08537155b' + .02494421c' &= .03263074,\\
.27700833a' + .07673361b' + .02125585c' &= .02705186,\\
.26298491a' + .06916105b' + .01818831c' &= .02231492,\\
.25000000a' + .06250000b' + .01562500c' &= .01831564,
\end{aligned}$$

where the second member is in each case the value of e^{-x^2}.

From these condition equations the normal equations are formed in the usual way, and are as follows:

$$\begin{aligned}
0.90288292a' + 0.29481990b' + 0.09812945c' &= 0.12839856\\
0.29481990a' + 0.09812945b' + 0.03324910c' &= 0.04358243\\
0.09812945a' + 0.03324910b' + 0.01145041c' &= 0.01502736
\end{aligned}$$

Solving these equations we find $\log a' = 9.2883208n$, $\log b' = 0.0713875$, $\log c' = 9.6489240n$; and from the known relations between a', b', c' and a, b, c, we have $\log a = 9.2883208$, $\log b = 9.5942662n$, $\log c = 8.9499540$, as given in the table. The constant A is determined from the condition that when $x = 1.6$ formula (2) should give the correct value of

$$\int_0^{1.6} e^{-x^2}\, dx.$$

A similar method was used in obtaining the other coefficients of the table;

but the number of condition equations was not always the same, and in some cases only 4 were used.

Theoretically it would seem as if a large increase in the number of condition equations would add largely to the accuracy of the results, but in practice I have not always found this to be the case.

In order to illustrate the use of the table let it be required to find the value of

$$\int_{0.45}^{1.61} e^{-x^2}\,dx.$$

The computation is as follows:

$$\begin{array}{rlrlrl}
\log a = & 9.2883208 & \log b = & 9.5942662n & \log c = & 8.9499540 \\
\log 1.61 = & \underline{0.2068259} & \log(1.61)^3 = & \underline{0.6204777} & \log(1.61)^5 = & \underline{1.0341295} \\
 & 9.0814949 & & 8.9737885n & & 7.9158245
\end{array}$$

$$\begin{array}{rl}
A = & 0.831292 \\
ax = & 0.120641 \\
cx^5 = & \underline{0.008238} \\
 & .960171 \\
bx^3 = & -\ \underline{.094143} \\
 & .866028 = \int_0^{1.61} e^{-x^2}\,dx
\end{array}$$

$$\begin{array}{rlrlrl}
\log a = & 9.9995025 & \log b = & 9.5148674n & \log c = & 8.9090072 \\
\log .45 = & \underline{9.6532125} & \log(.45)^3 = & \underline{8.9596375} & \log(.45)^5 = & \underline{8.2660625} \\
 & 9.6527150 & & 8.4745049n & & 7.1750697
\end{array}$$

$$\begin{array}{rl}
A = & 0.000225 \\
ax = & 0.449485 \\
cx^5 = & \underline{0.001496} \\
 & 0.451206 \\
bx^3 = & \underline{0.029820} \\
 & .421386 = \int_0^{.45} e^{-x^2}\,dx
\end{array}
\qquad
\begin{array}{l}
.866028 \\
\underline{.421386} \\
.444642 = \int_{0.45}^{1.61} e^{-x^2}\,dx.
\end{array}$$

If we attempt to compute

$$\int_0^{1.61} e^{-x^2}\,dx$$

by the three series commonly used, namely,

$$x - \frac{x^3}{3} + \frac{1}{1.2}\frac{x^5}{5} - \frac{1}{1.2.3}\frac{x^7}{7} + \frac{1}{1.2.3.4}\frac{x^9}{9} - \cdots,$$

$$e^{-x^2}x\left[1 + \frac{2x^2}{3} + \frac{(2x^2)^2}{3.5} + \frac{(2x^2)^3}{3.5.7} + \frac{(2x^2)^4}{3.5.7.9} + \cdots\right],$$

$$\tfrac{1}{2}\sqrt{\pi} - \frac{e^{-x^2}}{2x}\left[1 - \frac{1}{2x^2} + \frac{1.3}{(2x^2)^2} - \frac{1.3.5}{(2x^2)^3} + \frac{1.3.5.7}{(2x^2)^4} - \cdots\right],$$

we shall find that it will require 9 terms in each of the first two series to give the integral correct to three figures, while it cannot be obtained correct to three figures by using the third series, no matter how many terms are taken.

As previously stated, it is believed that the formulas which have been given will, in general, give the integral correct to five figures, for the reason that I have made comparisons of integrals computed by these formulas, for values of x from 0 to 2, and at intervals of every 0.02, with the values as given in standard tables. In the hundred comparisons thus made I found no discrepancy greater than one in the fifth figure.

The reason why the method above explained should be expected to give accurate results was explained in an article on the ballistic problem, published in the Annals of Mathematics for April, 1905.

It was not deemed advisable to extend the table of coefficients for values of x greater than 3, as the third series commonly used is very convenient for values of x beyond that limit.

Boston, Mass.

ON THE SECOND THEOREM OF THE MEAN*

BY CHARLES N. HASKINS

1. Introduction. The proofs of the second theorem of the mean are, with one exception, based either on the lemma of Abel or on the process of integration by parts.† The exception is the proof of Netto‡ which, however, imposes unnecessary restrictions on the functions involved, and is, moreover, even under these restrictions, incomplete.§ In the present note the lacuna of Netto's proof is filled, and, by use of the fundamental theorem of Fourier's constants,‖ the restrictions on the vanishing of the functions involved are removed.

2. The Second Theorem of the Mean. In the interval $a \leqq x \leqq b$ let the function $f(x)$ be single-valued, bounded, and monotonic; and let the function $\phi(x)$ be single-valued, and, as well as $[\phi(x)]^2$, integrable.¶ Then the second theorem of the mean asserts that

$$\int_a^b f(x)\,\phi(x)\,dx = f(a+)\int_a^\xi \phi(x)\,dx + f(b-)\int_\xi^b \phi(x)\,dx, \qquad a \leqq \xi \leqq b. \tag{1}$$

It is worth noting that if $\phi(x) \geqq 0$ the second theorem of the mean is an obvious consequence of the maximum-minimum theorem, provided $f(a+)$ and $f(b-)$ are the limits of indetermination of $f(x)$ in (a, b), in which case, moreover, $f(x)$ is not restricted to being monotonic. The two theorems are in a certain sense complementary; in the maximum-minimum theorem the

---

* Presented to the American Mathematical Society, New York, Feb. 28, 1908.

† Cf. Pringsheim, *Münchener Sitzungsberichte*, vol. 30 (1900), pp. 209-233.

‡ Netto, *Zeitschrift für Mathematik und Physik*, vol. 40 (1895), p. 180.

§ Pringsheim, loc. cit., p. 232.

‖ Cf. de la Vallée-Poussin, *Annales de la société scientifique de Bruxelles*, vol. 17 (1892-3), p. 18. Hurwitz, *Mathematische Annalen*, vol. 57 (1903), p. 425.

¶ The integrability of $[\phi(x)]^2$ brings about also the integrability of $|\phi(x)|$. The theorem is true even when $|\phi(x)|$ is not integrable provided $f(x) \cdot \phi(x)$ is so. The condition that $[\phi(x)]^2$ be integrable is introduced in order that the fundamental theorem of Fourier's constants may be used.

sign of $\phi(x)$ is restricted, but $f(x)$ need not be monotonic; in the second theorem of the mean the restriction on the sign of $\phi(x)$ is removed, but, as a compensation, the function $f(x)$ must be monotonic, as is shown by examples in which the theorem fails when $f(x)$ is not monotonic.*

3. Netto's proof and its limitations. If we discard the trivial case in which $f(a+) = f(b-)$, the relation (1) can be reduced to the simpler form

$$(2) \qquad \int_a^b F(x)\,\phi(x)\,dx = \int_a^\xi \phi(x)\,dx, \qquad a \leqq \xi \leqq b;$$

where

$$F(x) \equiv \frac{f(b-) - f(x)}{f(b-) - f(a+)}, \qquad a < x < b,$$

$$F(a) = F(a+) = 1, \quad F(b) = F(b-) = 0.$$

That is, if there exist a ξ for which either (1) or (2) is true then for that same ξ the other is true.

Netto's proof now proceeds as follows:

let $$K(x) \equiv \int_a^x F(x)\,\phi(x)\,dx, \qquad L(x) \equiv \int_a^x \phi(x)\,dx;$$

and consider the roots $a_1, a_2, \cdots$ of the equation

$$\phi(x) = 0.$$

It follows from the first law of the mean that the relation (2) holds in the interval $a \leqq x \leqq a_1$, and Netto then proves that if the relation holds in any interval $a \leqq x \leqq A$ it is possible to assign a $\delta > 0$ such that the relation holds in the interval $a \leqq x \leqq A + \delta$. From this he concludes that the relation must hold throughout the entire interval $a \leqq x \leqq b$. Now, as Pringsheim† points

---

* Cf. Kronecker-Netto, *Vorlesungen über die Theorie der einfachen und vielfachen Integrale*, p. 63. Pierpont, *Theory of Functions of Real Variables*, vol. 1, p. 380.

These examples leave open the question as to whether the validity of the theorem does not depend on the fact that $f(a+)$ and $f(b-)$ are the limits of indetermination of $f(x)$, rather than on the fact that $f(x)$ is monotonic. A very simple modification of Kronecker's example shows, however, that the latter is the determining condition.

† Pringsheim, loc. cit., p. 232.

out, this conclusion involves the restriction that $\phi(x) = 0$ shall have a *first* root $a_1 > a$, and the assumption that the points $A + \delta$ have not a limiting point $B < b$.

4. Proof of the non-existence of a limiting point. Under the hypotheses concerning $f(x)$ and $\phi(x)$ the two integrals $K(x)$ and $L(x)$ are continuous functions of x in the interval $a \leqq x \leqq b$. The assertion that for every x in the interval $a \leqq x \leqq A$ there exists a ξ, such that

$$K(x) = L(\xi), \qquad a \leqq \xi \leqq x,$$

is therefore equivalent to the assertion that in any interval $a \leqq x \leqq \bar{x} \leqq A$ the minimum and the maximum values of the continuous function $K(x)$ are respectively not less than the minimum and not greater than the maximum value of the continuous function $L(x)$. In other words,

$$\min \{L(x)\} \leqq \min \{K(x)\} \leqq \max \{K(x)\} \leqq \max \{L(x)\}, \qquad a \leqq x \leqq \bar{x} \leqq A.$$

If, then, there exist a limiting point B such that if $a \leqq x < B$, there always exists a ξ for which

$$K(x) = L(\xi), \qquad a \leqq \xi \leqq x < B;$$

but such that there exists no β for which

$$K(B) = L(\beta), \qquad a \leqq \beta \leqq B;$$

it follows that either

$$\left.\begin{aligned} &K(B) < \min L(x) \\ \text{or} \qquad &K(B) > \max L(x) \end{aligned}\right\}, \qquad a \leqq x \leqq B.$$

This, however, is impossible. For if $K(B)$ is neither the maximum nor the minimum value of $K(x)$ in $a \leqq x \leqq B$ then there exists an $x^* < B$ such that $K(x^*) = K(B)$, and corresponding to this x^*, a ξ^* such that

$$K(x^*) = L(\xi^*), \qquad a \leqq \xi^* \leqq x^* < B.$$

Hence

$$K(B) = L(\xi^*), \qquad a \leqq \xi^* < B.$$

Assume, then, that $K(B) = \max \{K(x)\}$, $a \leqq x \leqq B$. Take a set of values $x_1 < x_2 < x_3 \cdots < B$ approaching B as a limit. By the hypothesis, to each x_i corresponds at least one ξ_i such that

$$K(x_i) = L(\xi_i), \qquad a \leqq \xi_i \leqq x_i < B.$$

From the continuity of $K(x)$ follows that

$$\lim_{i=\infty} L(\xi_i) = \lim_{i=\infty} K(x_i) = K(B).$$

Now either

$$\left.\begin{array}{lll} & (1) & \min \{L(x)\} \leqq K(B) \leqq \max \{L(x)\} \\ \text{or} & (2) & K(B) < \min \{L(x)\} \\ \text{or} & (3) & \max \{L(x)\} < K(B) \end{array}\right\}, \quad a \leqq x \leqq B.$$

But (2) and (3) are impossible. For if (2) were true we could find an x_i with corresponding ξ_i such that

$$L(\xi_i) = K(x_i) < K(B) < \min \{L(x)\}, \quad a \leqq \xi_i \leqq x_i < B;$$

and if (3) were true we could find an x_i with corresponding ξ_i such that

$$\max \{L(x)\} < K(x_i) = L(\xi_i) < K(B), \quad a \leqq \xi_i \leqq x_i < B;$$

both of which are impossible.

Hence $$\min \{L(x)\} \leqq K(B) \leqq \max \{L(x)\}, \quad a \leqq x \leqq B,$$

and hence there exists a β for which

$$K(B) = L(\beta), \quad a \leqq \beta \leqq B.$$

The case $K(B) = \min \{K(x)\}$, $a \leqq x \leqq B$, may be treated in the same way and with the same result. Hence:

There exists no limiting point B such that for all $x < B$ a ξ can be found for which

$$K(x) = L(\xi), \quad a \leqq \xi \leqq x < B,$$

but such that no ξ can be found for which

$$K(B) = L(\xi), \quad a \leqq \xi \leqq B.$$

5. Application of the fundamental theorem of Fourier's constants. For convenience we will now take $a = 0$, $b = 2\pi$, which may always be done by a linear transformation of the variable x. The assumptions

concerning the nature of $f(x)$ and $\phi(x)$ ensure the existence of the Fourier's constants

$$a_k = \frac{1}{\pi}\int_0^{2\pi} F(x)\cos kx\,dx, \qquad \alpha_k = \frac{1}{\pi}\int_0^{2\pi}\phi(x)\cos kx\,dx,$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} F(x)\sin kx\,dx, \qquad \beta_k = \frac{1}{\pi}\int_0^{2\pi}\phi(x)\sin kx\,dx,$$

$$k = 0,\ 1,\ 2,\ \cdots.$$

Let
$$K = K(2\pi) = \int_0^{2\pi} F(x)\phi(x)\,dx,$$

$$K_n(x) = \frac{\alpha_0}{2}\int_0^x F(x)\,dx + \sum_{k=1}^{n}\left(\alpha_k\int_0^x F(x)\cos kx\,dx + \beta_k\int_0^x F(x)\sin kx\,dx\right)$$

$$= \int_0^x F(x)\left\{\frac{\alpha_0}{2} + \sum_{k=1}^{n}(\alpha_k\cos kx + \beta_k\sin kx)\right\}dx,$$

$$L_n(x) = \frac{\alpha_0 x}{2} + \sum_{k=1}^{n}\left\{\frac{\alpha_k}{k}\sin kx + \frac{\beta_k}{k}(1-\cos kx)\right\}.$$

It has been shown without the use of the second theorem of the mean* that

$$K = \pi\left[\frac{a_0\alpha_0}{2} + \sum_{k=1}^{\infty}(a_k\alpha_k + b_k\beta_k)\right],$$

$$L(x) = \frac{\alpha_0 x}{2} + \sum_{k=1}^{\infty}\left\{\frac{\alpha_k}{k}\sin kx + \frac{\beta_k}{k}(1-\cos kx)\right\}.$$

Moreover, the first of these series converges absolutely and the second converges absolutely and uniformly to the continuous function $L(x)$ in the interval $0 \leqq x \leqq 2\pi$. Hence

$$\lim_{n=\infty} K_n(2\pi) = K,$$

$$\lim_{n=\infty} L_n(x) = L(x), \qquad 0 \leqq x \leqq 2\pi.$$

---

*Cf. de la Vallée-Poussin, loc. cit., p. 32.
Hurwitz, loc. cit., p. 438.
Bôcher, Annals of Mathematics, ser. 2, vol. 7 (1908), p. 107.

Now $$L_n'(x) \equiv \frac{dL_n(x)}{dx} = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + \beta_k \sin kx,$$

and

$$L_n(x) = \int_0^x L_n'(x)\,dx, \qquad K_n(x) = \int_0^x F(x)\,L_n'(x)\,dx.$$

$L_n'(x)$ is a continuous function with but a finite number of changes of sign in the interval $0 \leqq x \leqq 2\pi$. To it, therefore, Netto's proof applies directly, and we have

$$K_n(2\pi) = L_n(\xi_n), \qquad 0 \leqq \xi_n \leqq 2\pi.$$

Hence $$K = \lim_{n=\infty} K_n(2\pi) = \lim_{n=\infty} L_n(\xi_n).$$

Now the points ξ_n have at least one limiting point ξ. Let $\xi_{n_1}, \xi_{n_2}, \cdots \xi_{n_k}, \cdots$ be a sub-set of the set of points ξ_n converging to ξ as a limit. Then

$$\lim_{k=\infty} L_{n_k}(\xi_{n_k}) = \lim_{n=\infty} L_n(\xi_n),$$

and, because of the uniform convergence of $L_n(x)$ to the limiting function $L(x)$,

$$\lim_{k=\infty} L_{n_k}(\xi_{n_k}) = L(\xi).$$

Hence

$$L(\xi) = K, \qquad 0 \leqq \xi \leqq 2\pi;$$

in other words, there exists a ξ such that

$$\int_0^{2\pi} F(x)\phi(x)dx = \int_0^{\xi} \phi(x)dx, \qquad 0 \leqq \xi \leqq 2\pi,$$

from which we have at once the second theorem of the mean as stated in paragraph 2.

6. Application of the fundamental theorem of Fourier's constants to the proof by Abel's lemma. The method of paragraph 5 was there applied to complete the proof given by Netto. It is equally applicable to the proof by means of Abel's lemma. $L_n'(x)$ changes sign at most a finite number $m_n - 1$ of times in the interval $0 \leqq x \leqq 2\pi$. Let the points at which it changes sign be $x_2, x_3, \cdots x_{m_n}$, and let $0 < x_1 < x_2$. Then

by the first theorem of the mean

$$K_n(x_1) - K_n(0) = \lambda_1 \{ L_n(x_1) - L_n(0) \},$$
$$K_n(x_2) - K_n(x_1) = \lambda_2 \{ L_n(x_2) - L_n(x_1) \},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$K_n(x_{m_n}) - K_n(x_{m_n-1}) = \lambda_{m_n} \{ L_n(x_{m_n}) - L_n(x_{m_n-1}) \},$$
$$K_n(2\pi) - K_n(x_{m_n}) = \lambda_{m_n+1} \{ L_n(2\pi) - L_n(x_{m_n}) \};$$

where $$1 \geqq \lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_{m_n} \geqq \lambda_{m_n+1} \geqq 0,$$

and these λ's are numbers lying between the bounds of indetermination of $F(x)$ in the respective intervals $0 \leqq x \leqq x_1$, $x_1 \leqq x \leqq x_2, \cdots x_{m_n-1} \leqq x \leqq x_{m_n}$, $x_{m_n} \leqq x \leqq 2\pi$. Since $L_n(0) = K_n(0) = 0$ we have

$$K_n(2\pi) = (\lambda_1 - \lambda_2)\, L_n(x_1) + (\lambda_2 - \lambda_3)\, L_n(x_2) \cdots$$
$$+ (\lambda_{m_n} - \lambda_{m_n+1})\, L_n(x_{m_n}) + \lambda_{m_n+1}\, L_n(2\pi).$$

Let l_n, L_n be respectively the minimum and the maximum of the continuous function $L_n(x)$ in $0 \leqq x \leqq 2\pi$. Then, since $\lambda_i - \lambda_{i+1} \geqq 0$, $\lambda_{m_n+1} \geqq 0$,

$$l_n \cdot (\lambda_1 - \lambda_2 + \lambda_2 - \lambda_3 + \cdots + \lambda_{m_n} - \lambda_{m_n+1} + \lambda_{m_n+1}) \leqq K_n(2\pi)$$
$$\leqq L_n \cdot (\lambda_1 - \lambda_2 + \lambda_2 - \lambda_3 + \cdots + \lambda_{m_n} - \lambda_{m_n+1} + \lambda_{m_n+1}).$$

Hence
$$l_n \lambda_1 \leqq K_n(2\pi) \leqq \lambda_1 L_n.$$

Let now x_1 approach zero. Then

$$\lim_{x_1=0} \lambda_1 = 1.$$

Hence $$l_n \leqq K_n(2\pi) \leqq L_n,$$

and therefore there exists at least one value ξ_n, $0 \leqq \xi_n \leqq 2\pi$, such that for $x = \xi_n$ the continuous function $L_n(x)$ takes on the value $K_n(2\pi)$, i. e.,

$$K_n(2\pi) = L_n(\xi_n).$$

The proof from this point onward then proceeds as in paragraph 5.

University of Illinois,
May 1, 1908.

ANOTHER PROOF OF A THEOREM IN MULTIPLY PERFECT NUMBERS

By R. D. Carmichael

J. Westlund has shown * that the only multiply perfect numbers of multiplicity 3 and of the form $p_1^{a_1} p_2^{a_2} p_3$, where p_1, p_2, p_3 are different primes and $p_1 < p_2 < p_3$, are $2^3 \cdot 3 \cdot 5$ and $2^5 \cdot 5 \cdot 7$. I have shown† that there is none of multiplicity 3 and of the form $p_1^{a_1} p_2^{a_2} p_3^{a_3}$, where $a_3 > 1$ and p_1, p_2, p_3 are as before. It is proposed now to give a very simple demonstration of the latter theorem. For this purpose we require to use a theorem employed in my paper on "Multiply perfect numbers of four different primes."‡ This proposition is as follows:

If x is a positive integer greater than 1, $x^t - 1$ has a prime factor not dividing $x^u - 1 (u < t)$, except in the two cases $t = 2$, $x = 2^v - 1$, $v \geqq 2$; and $t = 6$, $x = 2$. Such prime factors of $x^t - 1$ are of the form $st + 1$, and evidently if t is odd and greater than 1 they are of the form $2st + 1$.

For the problem in consideration we must have §

$$3 = \frac{2^{a_1+1} - 1}{2^{a_1}} \cdot \frac{3^{a_2+1} - 1}{3^{a_2} \cdot 2} \cdot \frac{p_3^{a_3+1} - 1}{p_3^{a_3}(p_3 - 1)},$$

if such multiply perfect numbers exist. But, by the theorem above quoted, if $a_3 > 1$, $p_3^{a_3+1} - 1$ must have a prime factor greater than 3, since the exceptional cases are excluded by the conditions imposed upon p_3 and a_3. Therefore the equation cannot be satisfied. Hence,

There are no multiply perfect numbers of multiplicity 3 and of the form $p_1^{a_1} p_2^{a_2} p_3^{a_3}$, where $p_1 < p_2 < p_3$ are different primes and $a_3 > 1$.

---

* Annals of Mathematics, ser. 2, vol. 2, p. 172 (1901).

† Ibid., vol. 7, p. 153 (1906), and vol. 8, p. 49 (1906).

‡ Ibid., vol. 8, p. 151 (1907).

§ Ibid., vol. 7, p. 153 (1906).

A THEOREM CONCERNING EQUAL RATIOS

By J. L. Coolidge

The following simple theorem is not included in the common textbooks on the Theory of Numbers, but it is altogether likely that it has been published before. The writer would be grateful to any reader of The Annals who should tell him where a demonstration has already appeared.

Theorem: If a set of ratios between positive integers are equal to one another, then all are equal to the ratio of the greatest common divisor of the numerators to that of the denominators, and also to the ratio of the least common multiple of the numerators to that of the denominators.

To prove the theorem, let

$$\frac{a_1}{b_1} = \frac{a_2}{b_2} \cdots = \frac{a_n}{b_n} = \frac{p}{q}$$

where p and q are any two positive integers in the same ratio as the positive integers, a_i and b_i where $i = 1, 2, \cdots n$. We have then n equations

$$qa_i = pb_i.$$

The greatest common divisor of the first members of these equations must be equal to that of the second members, so that if a be the greatest common divisor of the a_i's and b that of the b_i's, then, since qa and pb are the greatest common divisors of the first and second members respectively, we have

$$qa = pb \qquad \text{or} \qquad \frac{a}{b} = \frac{p}{q}.$$

If, further, we put

$$a_i = r_i a \qquad b_i = s_i b$$

then, since

$$\frac{a_i}{b_i} = \frac{a}{b}$$

we have $r_i = s_i$.

Now if A and B are the least common multiples of the a_i's and b_i's respectively, and if R and S are the least common multiples of the r_i's and s_i's respectively, we have

$$A = Ra \qquad B = Sb$$

and since, for all values of i, $r_i = s_i$, $R = S$, and consequently

$$\frac{A}{B} = \frac{a}{b} = \frac{p}{q}$$

CAMBRIDGE, MASS.,
JUNE, 1908.

NOTE ON CERTAIN ITERATED AND MULTIPLE INTEGRALS

By Wallie Abraham Hurwitz

Dirichlet* stated and used the following formula for inversion of the order of integration:

$$\int_0^a dx \int_0^x \phi(x, y)\,dy = \int_0^a dy \int_y^a \phi(x, y)\,dx.$$

Volterra† makes use of a similar formula, which he calls "Dirichlet's principle." The truth of the theorem is obvious when the integrand is continuous within and on the boundary of the field of integration. The object of the present paper is to justify the theorem when the integrand is allowed to become infinite in certain ways on the boundary of the field; and to extend it to space of n dimensions. Part II, which contains the generalisation to n dimensions, is entirely independent of Part I, and includes its results as a special case; the method of treatment is, however, slightly different.

While the results here stated, at least in so far as they relate to the case of two dimensions, may be deduced from a general theorem of de la Vallée Poussin,‡ the simple character of the reasoning here used seems to justify an independent treatment.

I.

A form of statement slightly different from that of Dirichlet and Volterra will be used, in which the integrals appear as the generalisation of integrals mentioned by Schlömilch.§

Theorem. *Let $f(x, y)$ be continuous and limited in the region*

$$R:\quad 0 < x, \quad 0 < y, \quad x + y < 1;$$

* *Crelle's Journal*, vol. 17 (1837), p. 45.

† Cf., for example, *Annali di Matematica*, vol. 25 (1897), p. 142.

‡ *Cours d'analyse infinitésimale*, vol. 2 (1906), §§ 73, 77.

§ *Analytische Studien*, vol. 1 (1848), §19. Dirichlet, Liouville and other writers had also discussed special cases of the integral here considered. Cf. the concluding paragraph of this paper.

and let $0 < \lambda, \mu, \nu \leqq 1$; *then*

$$\int_0^1 dx \int_0^{1-x} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y) dy$$

$$= \int_0^1 dy \int_0^{1-y} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y) dx. \quad (1)$$

Let $|f(x, y)| \leqq M$. The integral

$$I = \int_0^{1-x} x^{\lambda-1} y^{\mu-1} (1-x-y)^{\nu-1} f(x, y) dy$$

is equivalent, by means of the change of variable $y = (1-x)s$, to

$$x^{\lambda-1}(1-x)^{\mu+\nu-1} \int_0^1 s^{\mu-1} (1-s)^{\nu-1} f[x, (1-x)s] ds.$$

The expression $x^{\lambda-1} (1-x)^{\mu+\nu-1}$, which now stands outside the integral sign, is continuous for $0 < x < 1$; the integral is readily seen to be uniformly convergent for $0 \leqq x \leqq 1$, and hence to define a continuous function, whose absolute value is less than $MB(\mu, \nu)$. Thus the integral I represents a continuous function for $0 < x < 1$; and

$$|I| \leqq MB(\mu, \nu) \, x^{\lambda-1} (1-x)^{\mu+\nu-1}.$$

Furthermore, I may be integrated between the limits 0 and 1 ; for if we substitute for it the above expression greater than $|I|$, we obtain the convergent integral

$$MB(\mu, \nu) \int_0^1 x^{\lambda-1} (1-x)^{\mu+\nu-1} dx.$$

Hence the iterated integral on the left of (1) is convergent. Since the two sides of (1) differ, essentially, only in having x and y interchanged, it follows that the iterated integral on the right of (1) also converges. It is evident that if we integrate over any intervals not reaching outside the intervals appearing in the integrals just considered, we shall still have convergent integrals.

The field of integration is here a triangle bounded by the coordinate axes and the line $x + y = 1$. Draw lines inside this triangle, near its sides and parallel to them :

$$x = \delta, \quad y = \delta, \quad x + y = 1 - \delta.$$

Within and on the boundary of the smaller triangle thus formed the integrand is continuous; integration may therefore be performed in either order with the same result:

$$\int_\delta^{1-2\delta} dx \int_\delta^{1-x-\delta} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y)\,dy$$
$$= \int_\delta^{1-2\delta} dy \int_\delta^{1-y-\delta} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y)\,dx. \quad (2)$$

We may write

$$\int_\delta^{1-2\delta} dx \int_\delta^{1-x-\delta} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y)\,dy$$
$$= \int_\delta^{1-2\delta} dx \int_0^{1-x} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y)\,dy - I_1 - I_2, \quad (3)$$

where

$$I_1 = \int_\delta^{1-2\delta} dx \int_0^{\delta} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y)\,dy$$

and

$$I_2 = \int_\delta^{1-2\delta} dx \int_{1-x-\delta}^{1-x} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y)\,dy.$$

When $y < \delta$, $(1-x-y)^{\nu-1} \leqq (1-x-\delta)^{\nu-1}$; hence

$$\left|\int_0^{\delta} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y)\,dy\right| \leqq Mx^{\lambda-1}(1-x-\delta)^{\nu-1}\int_0^{\delta} y^{\mu-1}\,dy$$
$$= \frac{M}{\mu}\delta^{\mu} x^{\lambda-1}(1-x-\delta)^{\nu-1};$$

therefore $|I_1| \leqq \displaystyle\int_\delta^{1-2\delta} dx \left|\int_0^{\delta} x^{\lambda-1} y^{\mu-1}(1-x-y)^{\nu-1} f(x, y)\,dy\right|$

$$\leqq \frac{M}{\mu}\delta^{\mu}\int_\delta^{1-2\delta} x^{\lambda-1}(1-x-\delta)^{\nu-1}\,dx$$
$$\leqq \frac{M}{\mu}\delta^{\mu}\int_0^{1-\delta} x^{\lambda-1}(1-x-\delta)^{\nu-1}\,dx$$
$$= \frac{M}{\mu} B(\lambda, \mu)\,\delta^{\mu}(1-\delta)^{\lambda+\nu-1}. \quad (4)$$

Again, when $y > 1 - x - \delta$, $y^{\mu-1} \leqq (1 - x - \delta)^{\mu-1}$; hence

$$\left|\int_{1-x-\delta}^{1-x} x^{\lambda-1} y^{\mu-1}(1 - x - y)^{\nu-1} f(x, y)\, dy\right|$$
$$\leqq Mx^{\lambda-1}(1 - x - \delta)^{\mu-1}\int_{1-x-\delta}^{1-x}(1 - x - y)^{\nu-1}\, dy$$
$$= \frac{M}{\nu}\delta^{\nu} x^{\lambda-1}(1 - x - \delta)^{\mu-1};$$

therefore $|I_2| \leqq \int_{\delta}^{1-2\delta} dx \left|\int_{1-x-\delta}^{1-x} x^{\lambda-1} y^{\mu-1}(1 - x - y)^{\nu-1} f(x, y) dy\right|$

$$\leqq \frac{M}{\nu}\delta^{\nu}\int_{\delta}^{1-2\delta} x^{\lambda-1}(1 - x - \delta)^{\mu-1}\, dx$$
$$\leqq \frac{M}{\nu}\delta^{\nu}\int_{0}^{1-\delta} x^{\lambda-1}(1 - x - \delta)^{\mu-1}\, dx$$
$$= \frac{M}{\nu} B(\lambda, \nu)\, \delta^{\nu}(1 - \delta)^{\lambda+\mu-1}. \qquad (5)$$

From (4) and (5) we see that

$$\lim_{\delta=0} I_1 = 0, \qquad \lim_{\delta=0} I_2 = 0;$$

and hence from (3) that the limit of the left-hand side of (2) is

$$\lim_{\delta=0}\int_{\delta}^{1-2\delta} dx\int_{0}^{1-x} x^{\lambda-1} y^{\mu-1}(1 - x - y)^{\nu-1} f(x, y) dy$$
$$= \int_0^1 dx\int_0^{1-x} x^{\lambda-1} y^{\mu-1}(1 - x - y)^{\nu-1} f(x, y) dy. \qquad (6)$$

The two sides of (2) are obtained from each other by exchanging the rôles of x and y; it follows that also the limit of the right-hand side of (2) is

$$\int_0^1 dy\int_0^{1-y} x^{\lambda-1} y^{\mu-1}(1 - x - y)^{\nu-1} f(x, y) dx. \qquad (7)$$

If we now allow δ to approach zero in (2), making use of the facts stated in (6) and (7), we obtain the theorem to be proved.*

* It is in fact true that the *double* integral over the region R converges and is equal to the iterated integrals. The proof of this theorem, which would require separate treatment according as $f(x, y)$ does or does not retain the same sign throughout R, is a special case of the more general considerations of Part II.

The form in which the theorem is used by Dirichlet, and by Volterra and other writers on integral equations, is obtainable from this result by means of the substitution $x' = b - (b - a)x$. $y' = a + (b - a)y$.
We find thus the

THEOREM. *Let $\psi(x, y)$ be continuous and limited within the triangle bounded by the lines $x = y$, $x = b$, $y = a$; and let $0 < \lambda, \mu, \nu \leqq 1$; then*

$$\int_a^b dx \int_a^x (b - x)^{\lambda-1}(y - a)^{\mu-1}(x - y)^{\nu-1}\psi(x, y)dy$$
$$= \int_a^b dy \int_y^b (b - x)^{\lambda-1}(y - a)^{\mu-1}(x - y)^{\nu-1}\psi(x, y)dx.$$

The preceding theorems may be stated for a wider class of functions. The equality of the results of iterated integration in the two orders in the interior triangle is sufficient for the validity of the proof as given.

II.

The preceding considerations are readily extended to the case of n variables. We have first the following

THEOREM. *Consider the n-dimensional region*

$$R: \begin{cases} 0 < x_i \; [i = 1, 2, \cdots n]; \\ x_1 + x_2 + \cdots + x_n < 1; \end{cases}$$

and the $(n - 1)$-dimensional region

$$\bar{R}: \begin{cases} 0 < x_i \; [i = 1, 2, \cdots n - 1]; \\ x_1 + x_2 + \cdots + x_{n-1} < 1. \end{cases}$$

Let $f(x_1, x_2, \cdots x_n)$ be continuous and limited in R, and let

$$0 < \lambda_1, \lambda_2, \cdots \lambda_n, \lambda \leqq 1;$$

then

$$\int_R x_1^{\lambda_1-1} x_2^{\lambda_2-1} \cdots x_n^{\lambda_n-1}(1 - x_1 - x_2 - \cdots - x_n)^{\lambda-1} f dS$$
$$= \int_{\bar{R}} d\bar{S} \int_0^{1-x_1-x_2-\cdots-x_{n-1}} x_1^{\lambda_1-1} x_2^{\lambda_2-1} \cdots x_n^{\lambda_n-1}(1 - x_1 - x_2 - \cdots - x_n)^{\lambda-1} f dx_n.$$

It should first be noticed that when $f(x_1, x_2, \cdots x_n) = 1$, the n-fold

integral is known to be convergent, and its value is

$$\frac{\Gamma(\lambda_1)\Gamma(\lambda_2)\cdots\Gamma(\lambda_n)\Gamma(\lambda)}{\Gamma(\lambda_1+\lambda_2+\cdots+\lambda_n+\lambda)}.^*$$

It will be convenient to write

$$x_1^{\lambda_1-1}x_2^{\lambda_2-1}\cdots x_n^{\lambda_n-1}(1-x_1-x_2-\cdots-x_n)^{\lambda-1}=\phi(x_1, x_2, \cdots x_n).$$

Let $|f| \leqq M$, and suppose at first that $f \geqq 0$. The integral

$$I=\int_0^{1-x_1-\cdots-x_{n-1}}\phi\, dx_n,$$

subjected to the change of variable $x_n=(1-x_1-\cdots-x_{n-1})s$, becomes

$$x_1^{\lambda_1-1}\cdots x_{n-1}^{\lambda_{n-1}-1}(1-x_1-\cdots-x_{n-1})^{\lambda+\lambda_n-1}$$

$$\int_0^1 s^{\lambda_n-1}(1-s)^{\lambda-1}f[x_1, \cdots x_{n-1}, (1-x_1-\cdots-x_{n-1})s]\,ds.$$

The integral which appears in the last expression converges uniformly throughout the region formed by $\bar{R}$ and its boundary, and therefore represents a continuous function in this region; the function defined by I is therefore of the same character in $\bar{R}$ as the given function ϕ in R.† In $\bar{R}$, I is continuous, and

$$|I| \leqq MB(\lambda, \lambda_n)\, x_1^{\lambda_1-1}\cdots x_{n-1}^{\lambda_{n-1}-1}(1-x_1-\cdots-x_{n-1})^{\lambda+\lambda_n-1}$$

Also we may integrate I over $\bar{R}$, since the integral

$$MB(\lambda, \lambda_n)\int_{\bar{R}} x_1^{\lambda_1-1}\cdots x_{n-1}^{\lambda_{n-1}-1}(1-x_1-\cdots-x_{n-1})^{\lambda+\lambda_n-1}d\bar{S}$$

(which belongs to the special type just mentioned) is convergent.

* Goursat, *Cours d'analyse*, vol. I, §150. The proof there given under the hypothesis that the exponents are positive (so that the integral is *proper*) holds without alteration if each exponent is greater than -1.

† If $\lambda+\lambda_n \leqq 1$, we take the function defined by the above integral with respect to s as the function corresponding to f, and $\lambda+\lambda_n$ as the number corresponding to λ; if $\lambda+\lambda_n > 1$, we take the integral multiplied by $(1-x_1-\cdots-x_{n-1})^{\lambda+\lambda_n-1}$ as the function corresponding to f, and 1 as the number corresponding to λ.

Consider now the smaller fields of integration :

$$R_\delta: \begin{cases} \delta < x_i \quad [i = 1, 2, \cdots n]; \\ x_1 + x_2 + \cdots + x_n < 1 - \delta; \end{cases}$$

$$\bar{R}_\delta: \begin{cases} \delta < x_i \quad [i = 1, 2, \cdots n-1]; \\ x_1 + x_2 + \cdots + x_{n-1} < 1 - 2\delta; \end{cases}$$

$$\bar{R}': \begin{cases} 0 < x_i \quad [i = 1, 2, \cdots n-1]; \\ x_1 + x_2 + \cdots + x_{n-1} < 1 - \delta. \end{cases}$$

Since the integrand is continuous within and on the boundary of R_δ,

$$\int_{R_\delta} \phi dS = \int_{\bar{R}_\delta} d\bar{S} \int_\delta^{1-x_1-\cdots-x_{n-1}-\delta} \phi dx_n. \tag{8}$$

As before, write

$$\int_{\bar{R}_\delta} d\bar{S} \int_\delta^{1-x_1-\cdots-x_{n-1}-\delta} \phi dx_n = \int_{\bar{R}_\delta} d\bar{S} \int_0^{1-x_1-\cdots-x_{n-1}} \phi dx_n - I_1 - I_2, \tag{9}$$

where
$$I_1 = \int_{\bar{R}_\delta} d\bar{S} \int_0^\delta \phi dx_n,$$

and
$$I_2 = \int_{\bar{R}_\delta} d\bar{S} \int_{1-x_1-\cdots-x_{n-1}-\delta}^{1-x^1-\cdots-x_{n-}} \phi dx_n.$$

When $x_n < \delta$, $(1 - x_1 - \cdots - x_n)^{\lambda-1} \leqq (1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda-1}$; hence

$$\int_0^\delta \phi dx_n \leqq M x_1^{\lambda_1-1} \cdots x_{n-1}^{\lambda_{n-1}-1}(1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda-1} \int_0^\delta x_n^{\lambda_n-1} dx_n$$

$$= \frac{M}{\lambda_n} \delta^{\lambda_n} x_1^{\lambda_1-1} \cdots x_{n-1}^{\lambda_{n-1}-1}(1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda-1};$$

therefore

$$I_1 \leqq \frac{M}{\lambda_n} \delta^{\lambda_n} \int_{\bar{R}_\delta} x_1^{\lambda_1 - 1} \cdots x_{n-1}^{\lambda_{n-1}-1}(1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda - 1} d\bar{S}$$

$$\leqq \frac{M}{\lambda_n} \delta^{\lambda_n} \int_{\bar{R}'} x_1^{\lambda_1 - 1} \cdots x_{n-1}^{\lambda_{n-1}-1}(1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda - 1} d\bar{S}^*$$

$$= \frac{M}{\lambda_n} \frac{\Gamma(\lambda_1) \cdots \Gamma(\lambda_{n-1})\Gamma(\lambda)}{\Gamma(\lambda_1 + \cdots + \lambda_{n-1} + \lambda)} \delta^{\lambda_n} (1 - \delta)^{\lambda_1 + \cdots + \lambda_{n-1} + \lambda - 1} \qquad (10)$$

Also, when

$$x_n > (1 - x_1 - \cdots - x_{n-1} - \delta), \; x_n^{\lambda_n - 1} \leqq (1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda_n - 1};$$

hence $\displaystyle\int_{1-x_1-\cdots-x_{n-1}-\delta}^{1-x_1-\cdots-x_{n-1}} \phi\, dx_n$

$$\leqq M x_1^{\lambda_1 - 1} \cdots x_{n-1}^{\lambda_{n-1}-1}(1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda_n - 1} \int_{1-x_1-\cdots-x_{n-1}-\delta}^{1-x_1-\cdots-x_{n-1}} (1 - x_1 - \cdots - x_n)^{\lambda - 1} dx_n$$

$$= \frac{M}{\lambda} \delta^{\lambda} x_1^{\lambda_1 - 1} \cdots x_{n-1}^{\lambda_{n-1}-1}(1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda_n - 1};$$

therefore

$$I_2 \leqq \frac{M}{\lambda} \delta^{\lambda} \int_{\bar{R}_\delta} x_1^{\lambda_1 - 1} \cdots x_{n-1}^{\lambda_{n-1}-1}(1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda_n - 1} d\bar{S}$$

$$\leqq \frac{M}{\lambda} \delta^{\lambda} \int_{\bar{R}'} x_1^{\lambda_1 - 1} \cdots x_{n-1}^{\lambda_{n-1}-1}(1 - x_1 - \cdots - x_{n-1} - \delta)^{\lambda_n - 1} d\bar{S}$$

$$= \frac{M}{\lambda} \frac{\Gamma(\lambda_1) \cdots \Gamma(\lambda_{n-1})\Gamma(\lambda_n)}{\Gamma(\lambda_1 + \cdots + \lambda_{n-1} + \lambda_n)} \delta^{\lambda}(1 - \delta)^{\lambda_1 + \cdots + \lambda_{n-1} + \lambda_n}. \qquad (11)$$

From (10) and (11) we see that

$$\lim_{\delta = 0} I_1 = 0, \qquad \lim_{\delta = 0} I_2 = 0;$$

and therefore from (9) that the limit of the right-hand side of (8) is

$$\lim_{\delta=0} \int_{\bar{R}_\delta} d\bar{S} \int_0^{1-x_1-\cdots-x_{n-1}} \phi\, dx_n = \int_{\bar{R}} d\bar{S} \int_0^{1-x_1-\cdots-x_{n-1}} \phi\, dx_n. \qquad (12)$$

---

* This integral is evaluated by means of the substitution $x_i = (1 - \delta)y_i$, $[i = 1, 2, \cdots n - 1]$ which reduces it to the special case already mentioned.

As for the left-hand side of (8), we note that since the integrand *retains the same sign throughout* R and since for a *single set of regions* R_δ whose limit is R,

$$\lim_{\delta=0}\int_{R_\delta}\phi ds$$

exists, therefore the integral over R converges and

$$\int_R \phi dS = \lim_{\delta=0}\int_{R_\delta}\phi dS.$$

Hence we conclude from (8) and (12) that

$$\int_R \phi dS = \int_{\bar{R}} d\bar{S}\int_0^{1-x_1-\cdots-x_{n-1}}\phi dx_n,$$

which was the theorem to be proved.

Suppose now that f changes sign in R. It will always be possible to express f as the difference of two functions which are nowhere negative, each of which is continuous and limited in R.* The theorem will hold for each of these functions, and hence for their difference f.

We have thus reduced an n-fold integral to an $(n-1)$-fold integral of a simple integral. In the course of the proof it appeared that the function of $n-1$ variables resulting from the simple integration has the same properties in $\bar{R}$ as the given function of n variables has in R. We may therefore repeat the process time after time, until we obtain an iterated integral in n variables. Furthermore, at each step the selection of the variable with respect to which the simple integration is performed is evidently only a matter of notation; we have therefore the

THEOREM. *Let* $f(x_1, x_2, \cdots x_n)$ *be continuous and limited in the region*

$$R:\begin{cases} 0 < x_i \quad [i = 1, 2, \cdots n]; \\ x_1 + x_2 + \cdots + x_n < 1; \end{cases}$$

and let $0 < \lambda_1, \lambda_2 \cdots \lambda_n \leqq 1$; *then the* n*-fold integral*

$$\int_R x_1^{\lambda_1-1}x_2^{\lambda_2-1}\cdots x_n^{\lambda_n-1}(1-x_1-x_2-\cdots-x_n)^{\lambda-1}fdS$$

* For instance, by writing $f = M - (M - f)$, where M is a constant greater than any value of f in R.

converges, and may be evaluated by iterated integration in any order; for example,

$$\int_R x_1^{\lambda_1-1} x_2^{\lambda_2-1} \cdots x_n^{\lambda_n-1}(1 - x_1 - x_2 - \cdots - x_n)^{\lambda-1} f\,dS$$

$$= \int_0^1 x_1^{\lambda_1-1}\,dx_1 \int_0^{1-x_1} x_2^{\lambda_2-1}\,dx_2 \cdots \int_0^{1-x_1-x_2-\cdots-x_{n-1}} x_n^{\lambda_n-1}(1 - x_1 - x_2 - \cdots - x_n)^{\lambda-1} f\,dx_n.$$

By a substitution of the form $x_i = \left(\frac{x_i'}{a_i}\right)^{\rho_i}$ and a slight change of notation, the theorem may be made to apply to the more general case in which the region of integration is

$$\begin{cases} 0 < x_i & [i = 1, 2, \cdots n]; \\ \left(\frac{x_1}{a_1}\right)^{\rho_1} + \left(\frac{x_2}{a_2}\right)^{\rho_2} + \cdots + \left(\frac{x_n}{a_n}\right)^{\rho_n} < 1; \end{cases}$$

and the integrand is

$$x_1^{\lambda_1-1} x_2^{\lambda_2-1} \cdots x_n^{\lambda_n-1}\left[1 - \left(\frac{x_1}{a_1}\right)^{\rho_1} - \left(\frac{x_2}{a_2}\right)^{\rho_2} - \cdots - \left(\frac{x_n}{a_n}\right)^{\rho_n}\right]^{\lambda-1} f(x_1, x_2, \cdots x_n),$$

where $0 < \lambda_i \leqq \rho_i\ [i = 1, 2, \cdots n]$, $\quad 0 < \lambda \leqq 1$.

Special cases of integrals of the types discussed in this paper were considered by Dirichlet,* and shortly afterward by Liouville † and Catalan.‡ The most general case (belonging to this type) mentioned by these writers is obtained by taking for $f(x_1, x_2, \cdots x_n)$ in this paper a function of $x_1 + x_2 + \cdots + x_n$. The problem discussed is always the evaluation of the integral, or at least its reduction to a simple integral, —all necessary considerations as to convergence and related questions being assumed. All these results are given by Schlömilch.§ References to later work of like character will be found in *Encyclopädie der Mathematischen Wissenschaften*, II A3, footnotes 147–149.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.,
MARCH, 1908.

* *Abhandlungen der Akademie der Wissenschaften zu Berlin*, 1839, p. 61.
† *Journal de Mathématiques*, vol. 4 (1839), p. 225.
‡ *Journal de Mathématiques*, vol. 4 (1839), p. 323.
§ Loc. cit.

OCTOBER 1907

ANNALS OF MATHEMATICS

(Founded by Ormond Stone)

PUBLISHED UNDER THE AUSPICES OF HARVARD UNIVERSITY

Second Series Vol. 9 No. 1

FOR SALE BY
THE PUBLICATION OFFICE OF HARVARD UNIVERSITY
2 University Hall, Cambridge, Mass., U. S. A.

London: Longmans, Green & Co.
39 Paternoster Row

Leipzig: Otto Harrassowitz
Querstrasse 14

The following pamphlets, most of which have been reprinted from the ANNALS OF MATHEMATICS, second series, will be sent post-paid on receipt of price by the PUBLICATION OFFICE OF HARVARD UNIVERSITY, 2 University Hall, Cambridge, Mass.

CONTENTS

ANNALS OF MATHEMATICS

Published in October, January, April, and July, under the auspices of Harvard University, Cambridge, Mass., U. S. A.

Cambridge: Address *The Annals of Mathematics*, 2 University Hall, Cambridge, Mass., U. S. A. Subscription price, $2 a volume (four numbers) in advance. Single numbers, 75c. All drafts and money orders should be made payable to Harvard University.

London: Longmans, Green & Co., 39 Paternoster Row. Price, 2 shillings a number.

Leipzig: Otto Harrassowitz, Querstrasse 14. Price, 2 marks a number.

PRINTED BY THE SALEM PRESS CO., SALEM, MASS., U. S. A.

JULY 1908

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

PUBLISHED UNDER THE AUSPICES OF HARVARD UNIVERSITY

SECOND SERIES Vol. 9 No. 4

FOR SALE BY
THE PUBLICATION OFFICE OF HARVARD UNIVERSITY
2 University Hall, Cambridge, Mass., U. S. A.

London: LONGMANS, GREEN & Co.
39 Paternoster Row

Leipzig: OTTO HARRASSOWITZ
Querstrasse 14

CONTENTS

ANNALS OF MATHEMATICS

Published in October, January, April, and July, under the auspices of Harvard University, Cambridge, Mass., U. S. A.

Cambridge: Address *The Annals of Mathematics*, 2 University Hall, Cambridge, Mass., U. S. A. Subscription price, $2 a volume (four numbers) in advance. Single numbers, 75c. All drafts and money orders should be made payable to Harvard University.

London: Longmans, Green & Co., 39 Paternoster Row. Price, 2 shillings a number.

Leipzig: Otto Harrassowitz, Querstrasse 14. Price, 2 marks a number.

PRINTED BY THE SALEM PRESS CO., SALEM, MASS., U. S. A.